FIGURE 1

CCAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT

CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA

CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC

AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC

TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGGAGGCACAGGTGGCCCCCACCACCCGGAGG

AGCAGCTCCTGCCCCTGTCCGGGGG**ATG**ACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGA

AGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCT

TCTGGTGTTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTTAGGGTGTGTGCT

GTCCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTCGTGCAGCGTGTGTACCAGCCCTTCC

TCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATTTATAGGACCGCCTAC

CGCCGCAGCCCTGGGCTGGCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAG

GACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAG

GGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAG

TCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGC

CGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTG

TGCCCAAGGGAGGGCCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAG

GAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCT

GGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCC

TCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCTTC

CTGGAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGACTCG**TGA**CTGCCCAGCGCCCCAGG

CTGGACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCCTGCCCAACATGCTGGGGGTC

CAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTTTTCCTCCTC

CCCTTCCCTCGGGAGGGTCCCCAGACCCTGGCATGGGATGGGCTGGGATTTTTTTTGTGAAT

CCACCCCTGGCTACCCCCACCCTGGTTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCA

GCTGAGGGAAGGTACGAGTTCCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCC

CGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAA

CGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCT

AGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGT

TACAAAT

FIGURE 2

MTDSPPPGHPEEKATPPGGTGHEGLSGGAADVASGVGSGRHRARLPARPLGCVLSRAHGDPV

SESFVQRVYQPFLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC

GAAICQPPCRNGGSCVQPGRCRCPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCW

EGHSLSADGTLCVPKGGPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLAS

QALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 93-97, 270-274

N-myristoylation sites.

amino acids 19-25, 78-84, 97-103, 100-106, 103-109, 157-163, 191-197, 265-271

Amidation site.

amino acids 26-30

Aspartic acid and asparagine hydroxylation site.

amino acids 152-164

Cell attachment sequence.

amino acids 130-133

EGF-like domain cysteine pattern signature.

amino acids 123-135

FIGURE 3

CGCTCGCCCCGTCGCCCCTCGCCTCCCCGCAGAGTCCCCTCGCGGCAGCAGATGTGTGTGGG
GTCAGCCCACGGCGGGGACT**ATG**GTGAAATTCCCGGCGCTCACGCACTACTGGCCCCTGATC
CGGTTCTTGGTGCCCCTGGGCATCACCAACATAGCCATCGACTTCGGGGAGCAGGCCTTGAA
CCGGGGCATTGCTGCTGTCAAGGAGGATGCAGTCGAGATGCTGGCCAGCTACGGGCTGGCGT
ACTCCCTCATGAAGTTCTTCACGGGTCCCATGAGTGACTTCAAAAATGTGGGCCTGGTGTTT
GTGAACAGCAAGAGAGACAGGACCAAAGCCGTCCTGTGTATGGTGGTGGCAGGGGCCATCGC
TGCCGTCTTTCACACACTGATAGCTTATAGTGATTTAGGATACTACATTATCAATAAACTGC
ACCATGTGGACGAGTCGGTGGGGAGCAAGACGAGAAGGGCCTTCCTGTACCTCGCCGCCTTT
CCTTTCATGGACGCAATGGCATGGACCCATGCTGGCATTCTCTTAAAACACAAATACAGTTT
CCTGGTGGGATGTGCCTCAATCTCAGATGTCATAGCTCAGGTTGTTTTTGTAGCCATTTTGC
TTCACAGTCACCTGGAATGCCGGGAGCCCCTGCTCATCCCGATCCTCTCCTTGTACATGGGC
GCACTTGTGCGCTGCACCACCCTGTGCCTGGGCTACTACAAGAACATTCACGACATCATCCC
TGACAGAAGTGGCCCGGAGCTGGGGGGAGATGCAACAATAAGAAAGATGCTGAGCTTCTGGT
GGCCTTTGGCTCTAATTCTGGCCACACAGAGAATCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTCACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTCG
ACAAGAATAACCCCAGCAACAAACTGGTGAGCACGAGCAACACAGTCACGGCAGCCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGAGAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTTGTTCCTTTGCGGATCTTCTCCTTCTTCCCAGTTCCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTCCTTGCCCCCAGCTCTGTGCTGCG
GATCATCGTCCTCATCGCCAGCCTCGTGGTCCTACCCTACCTGGGGGTGCACGGTGCGACCC
TGGGCGTGGGCTCCCTCCTGGCGGGCTTTGTGGGAGAATCCACCATGGTCGCCATCGCTGCG
TGCTATGTCTACCGGAAGCAGAAAAAGAAGATGGAGAATGAGTCGGCCACGGAGGGGGAAGA
CTCTGCCATGACAGACATGCCTCCGACAGAGGAGGTGACAGACATCGTGGAAATGAGAGAGG
AGAATGAA**TAA**GGCACGGGACGCCATGGGCACTGCAGGGACGGTCAGTCAGGATGACACTTC
GGCATCATCTCTTCCCTCTCCCATCGTATTTTGTTCCCTTTTTTTTGTTTTGTTTTGGTAAT
GAAAGAGGCCTTGATTTAAAGGTTTCGTGTCAATTCTCTAGCATACTGGGTATGCTCACACT
GACGGGGGGACCTAGTGAATGGTCTTTACTGTTGCTATGTAAAAACAAACGAAACAACTGAC
TTCATACCCCTGCCTCACGAAAACCCAAAAGACACAGCTGCCTCACGGTTGACGTTGTGTCC
TCCTCCCCTGGACAATCTCCTCTTGGAACCAAAGGACTGCAGCTGTGCCATCGCGCCTCGGT
CACCCTGCACAGCAGGCCACAGACTCTCCTGTCCCCCTTCATCGCTCTTAAGAATCAACAGG
TTAAAACTCGGCTTCCTTTGATTTGCTTCCCAGTCACATGGCCGTACAAAGAGATGGAGCCC
CGGTGGCCTCTTAAATTTCCCTTCTGCCACGGAGTTCGAAACCATCTACTCCACACATGCAG
GAGGCGGGTGGCACGCTGCAGCCCGGAGTCCCCGTTCACACTGAGGAACGGAGACCTGTGAC
CACAGCAGGCTGACAGATGGACAGAATCTCCCGTAGAAAGGTTTGGTTTGAAATGCCCCGGG
GGCAGCAAACTGACATGGTTGAATGATAGCATTTCACTCTGCGTTCTCCTAGATCTGAGCAA
GCTGTCAGTTCTCACCCCCACCGTGTATATACATGAGCTAACTTTTTTAAATTGTCACAAAA
GCGCATCTCCAGATTCCAGACCCTGCCGCATGACTTTTCCTGAAGGCTTGCTTTTCCCTCGC
CTTTCCTGAAGGTCGCATTAGAGCGAGTCACATGGAGCATCCTAACTTTGCATTTTAGTTTT
TACAGTGAACTGAAGCTTTAAGTCTCATCCAGCATTCTAATGCCAGGTTGCTGTAGGGTAAC
TTTTGAAGTAGATATATTACCTGGTTCTGCTATCCTTAGTCATAACTCTGCGGTACAGGTAA
TTGAGAATGTACTACGGTACTTCCCTCCCACCATACGATAAAGCAAGACATTTTATAACG
ATACCAGAGTCACTATGTGGTCCTCCCTGAAATAACGCATTCGAAATCCATGCAGTGCAGTA
TATTTTTCTAAGTTTTGGAAAGCAGGTTTTTTCCTTTAAAAAAATTATAGACACGGTTCACT
AAATTGATTTAGTCAGAATTCCTAGACTGAAGAACCTAAACAAAAAAATATTTTAAAGATA
TAAATATATGCTGTATATGTTATGTAATTTATTTTAGGCTATAATACATTTCCTATTTTCGC
ATTTTCAATAAAATGTCTCTAATACAAAAAA

FIGURE 4

MVKFPALTHYWPLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLAYSLMKFF

TGPMSDFKNVGLVFVNSKRDRTKAVLCMVVAGAIAAVFHTLIAYSDLGYYIINKLHHVDESV

GSKTRRAFLYLAAFPFMDAMAWTHAGILLKHKYSFLVGCASISDVIAQVVFVAILLHSHLEC

REPLLIPILSLYMGALVRCTTLCLGYYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL

ATQRISRPIVNLFVSRDLGGSSAATEAVAILTATYPVGHMPYGWLTEIRAVYPAFDKNNPSN

KLVSTSNTVTAAHIKKFTFVCMALSLTLCFVMFWTPNVSEKILIDIIGVDFAFAELCVVPLR

IFSFFPVPVTVRAHLTGWLMTLKKTFVLAPSSVLRIIVLIASLVVLPYLGVHGATLGVGSLL

AGFVGESTMVAIAACYVYRKQKKKMENESATEGEDSAMTDMPPTEEVTDIVEMREENE

Transmembrane domains:

amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374,
408-423, 431-445

FIGURE 5

CCTGACAGAAGTGCCCCGGAGCTGGGGGAGATNCAACATTAAGAAGATGCTGAGCTTCTGGT

GCCNTTTGGCTCTAATTCTGGCCACACAGAGAANCAGTCGGCCTATTGTCAACCTCTTTGTT

TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA

CCCTGTGGGTCACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTCG

ACAAGAATAACCCCAGCAACAAACTGGTGAGCACGAGCAACACAGTCACGGCGGCCCACATC

AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC

ACCCAACGTGTCTGNGAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC

TCTGTGTTGTTCCTTTGCGGATCTTCTCCTTCTTCCCAGTTCCAGTCACAGTGAGGGCGCAT

CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTC

FIGURE 6

```
TGACGGAATCCCGGGCTGGGTATCCTGGTTTNGACAAGATAAACCCCCAGCAANAAATTGGG
GAGCAGGGCAAAACAGTNACGGGCAGCCCACATCAAGAAGTTCACCTTNGTTTGNATGGNTC
TGTCAACTCACGCTNTGTTTCGTGATGTTTTGGACACCCAAAGTGTTTGAGAAAATTTTGAT
AGACATNATCGGAGTGGANTTTGCCTTTGCAGAANTTTGNGNTGTTCCTTTGCGGATTTTCT
CCTTTTTCCCAGTTCCAGTCACAGNGAGGGCGCATCTCACCGGGNGGNTGATGACANTGAAG
AAAACCTTTGTCCTTGCCCCCAGCTNTTTGGTGCGGATCATTGTCCTNATNGCCAGCCTTGT
GGTCCTACCCTACCTGGGGGTGCACGGTGCGACCCTGGGCGTGGGTTCCCTCCTGGCGGGCA
```

FIGURE 7

TATTCCCAGTTCCGGTCACGGGGAGGGCGCATNTCACCGGGTGGCTGANGACACTGAAGAAA

ACCTTNGTCCTTGCCCCCAGNTTTGTGNTGCGGATNATCGTCCTCATCGCCAGCCTNGTGGT

CCTACCCTACCTGGGGGTGCACGGTGAGAC

FIGURE 8

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCC**ATG**GGGGCCTGCCTGGGA

GCCTGCTCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAG

CTGCTGCCCCGCCAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCCTCTTCC

TGGGGGTGCTGGTGTCCATCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTG

CCCTGGGTGTGTGAGGAGGGGGCCGGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGG

CTCCCTGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCT

TCTTTTTCACCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAG

AATGGGTTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACAT

CCCTGACGGCTCCTTCACCAACATCTGGTTCTACTTCGGCGTCGTGGGCTCCTTCCTCTTCA

TCCTCATCCAGCTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGC

AAGGCCGAGGAGTGCGATTCCCGTGCCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTCTT

CTACTTGCTGTCGATCGCGGCCGTGGCGCTGATGTTCATGTACTACACTGAGCCCAGCGGCT

GCCACGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCT

GCTGTCCTGCCCAAGGTCCAGGACGCCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTCAT

CACCCTCTACACCATGTTTGTCACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCA

ACCCCCATTTGCCAACCCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAG

ACCCAGTGGTGGGATGCCCCGAGCATTGTGGGCCTCATCATCTTCCTCCTGTGCACCCTCTT

CATCAGTCTGCGCTCCTCAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCC

CACCTATGCTAGACGCCACACAGCAGCAGCAGCAGCAGGTGGCAGCCTGTGAGGGCCGGGCC

TTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGCT

GGCCTCACTGCACGTCATGATGACGCTCACCAACTGGTACAAGCCCGGTGAGACCCGGAAGA

TGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTC

TACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGC**TGA**GGCAGCCT

CACAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCT

GCCCCCTCCCCACACCAATCAGCCAGGCTGAGCCCCCACCCCTGCCCCAGCTCCAGGACCTG

CCCCTGAGCCGGGCCTTCTAGTCGTAGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCA

GAGCCCCATCCCCCCGCCACACCCACACGGTGGAGCTGCCTCTTCCTTCCCCTCCTCCCTGT

TGCCCATACTCAGCATCTCGGATGAAAGGGCTCCCTTGTCCTCAGGCTCCACGGGAGCGGGG

CTGCTGGAGAGAGCGGGGAACTCCCACCACAGTGGGGCATCCGGCACTGAAGCCCTGGTGTT

CCTGGTCACGTCCCCCAGGGGACCCTGCCCCCTTCCTGGACTTCGTGCCTTACTGAGTCTCT

AAGACTTTTTCTAATAAACAAGCCAGTGCGTGTAAAAAAAA

FIGURE 9

MGACLGACSLLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVE

SQLYKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSSRD

PRAAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSW

NQRWLGKAEECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFC

VCVSIAAVLPKVQDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVA

GPEGYETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVA

ACEGRAFDNEQDGVTYSYSFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVWVKICAS

WAGLLLYLWTLVAPLLLRNRDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 10

GAGCGAGGCCGGGGACTGAAGGTGTGGGTGTCGAGCCCTCTGGCAGAGGGTTAACCTGGGTC

AAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCCGCGGCACGTCCGCGAGGACTTGA

AGTCCTGAGCGCTCAAGTTTGTCCGTAGGTCGAGAGAAGGCC**ATG**GAGGTGCCGCCACCGGC

ACCGCGGAGCTTTCTCTGTAGAGCATTGTGCCTATTTCCCCGAGTCTTTGCTGCCGAAGCTG

TGACTGCCGATTCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCC

TATTACCCGGAATCTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAG

AATTTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGG

TGTATGGGGGAATACCAGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCA

GAAATTTATCATAACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTT

CATTCGTTATGGCTGGCGCTGGGGTTGGAGAACTGCAGTGTTTGTGACTATATTCAACACAG

TGAACACTAGTCTGAATGTATACCGAAATAAAGATGCCTTAAGCCATTTTGTAATTGCAGGA

GCTGTCACGGGAAGTCTTTTTAGGATAAACGTAGGCCTGCGTGGCCTGGTGGCTGGTGGCAT

AATTGGAGCCTTGCTGGGCACTCCTGTAGGAGGCCTGCTGATGGCATTTCAGAAGTACGCTG

GTGAGACTGTTCAGGAAAGAAAACAGAAGGATCGAAAGGCACTCCATGAGCTAAAACTGGAA

GAGTGGAAAGGCAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTTACG

GGAAGATGAACCTGAGAATGATGCTAAGAAAATTGAAGCACTGCTAAACCTTCCTAGAAACC

CTTCAGTAATAGATAAACAAGACAAGGAC**TGA**AAGTGCTCTGAACTTGAAACTCACTGGAGA

GCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTTGGTCAGCCTGC

TGACAAATTTAAGTGCTGGTACCTGTGGTGGCAGTGGCTTGCTCTTGTCTTTTTCTTTTCTT

TTTAACTAAGAATGGGGCTGTTGTACTCTCACTTTACTTATCCTTAAATTTAAATACATACT

TATGTTTGTATTAATCTATCAATATATGCATACATGGATATATCCACCCACCTAGATTTTAA

GCAGTAAATAAACATTTCGCAAAGATTAAAGTTGAATTTTACAGTTT

FIGURE 11

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318

><subunit 1 of 1, 285 aa, 1 stop

><MW: 32190, pI: 9.03, NX(S/T): 2

MEVPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRLRELF

GKDEQQRISKDLANICKTAATAGIIGWVYGGIPAFIHAKQQYIEQSQAEIYHNRFDAVQSAH

RAATRGFIRYGWRWGWRTAVFVTIFNTVNTSLNVYRNKDALSHFVIAGAVTGSLFRINVGLR

GLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEEWKGRLQVTEHLPE

KIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and 171-195

N-glycosylation site:

amino acids 153-156

FIGURE 12

CGGAAGTCCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA

ATCTGGATGGGACCGCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGGA

CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGAA

TACCAGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATCAT

AACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTTCATTCGTTCATG

GCTGGCGCCGAACC

FIGURE 13

TCAAGTTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGCACCGCGGAGCTT
TTTTCTGTAGAGCATTGTGCCTATTTCCCCGAGTTTTTGCTGCCGAAGCTGTGACTGCCGAT
TCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATTTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGG
ACCTTGCTGATATNTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGA
ATACCAGCTTTTATTCATGNTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATNA
TAACC

FIGURE 14

GAGCCGCCGCCGCGCGCGCCGCGCACTGCAGCCCCAGGCCCCGGCCCCCCACCCACGTCT

GCGTTGCTGCCCCGCCTGGGCCAGGCCCCAAAGGCAAGGACAAAGCAGCTGTCAGGGAACCT

CCGCCGGAGTCGAATTTACGTGCAGCTGCCGGCAACCACAGGTTCCAAG**ATG**GTTTGCGGGG

GCTTCGCGTGTTCCAAGAACTGCCTGTGCGCCCTCAACCTGCTTTACACCTTGGTTAGTCTG

CTGCTAATTGGAATTGCTGCGTGGGGCATTGGCTTCGGGCTGATTTCCAGTCTCCGAGTGGT

CGGCGTGGTCATTGCAGTGGGCATCTTCTTGTTCCTGATTGCTTTAGTGGGTCTGATTGGAG

CTGTAAACATCATCAGGTGTTGCTATTTTTTTTATATGATTATTCTGTTACTTGTATTTATT

GTTCAGTTTTCTGTATCTTGCGCTTGTTTAGCCCTGAACCAGGAGCAACAGGGTCAGCTTCT

GGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAATCTAAACTGCT

GTGGGTTCCGAAGTGTTAACCCAAATGACACCTGTCTGGCTAGCTGTGTTAAAAGTGACCAC

TCGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGAGATTTGTTGG

TGGCATTGGCCTGTTCTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACCTACAGATACA

GGAACCAGAAAGACCCCGCGCGAATCCTAGTGCATTCCTT**TGA**TGAGAAACAAGGAAGAT

TTCCTTTCGTATTATGATCTTGTTCACTTTCTGTAATTTTCTGTTAAGCTCCATTTGCCAGT

TTAAGGAAGGAAACACTATCTGGAAAGTACCTTATTGATAGTGGAATTATATATTTTTACT

CTATGTTTCTCTACATGTTTTTTTTCTTTCCGTTGCTGAAAAATATTTGAAACTTGTGGTCTC

TGAAGCTCGGTGGCACCTGGAATTTACTGTATTCATTGTCGGGCACTGTCCACTGTGGCCTT

TCTTAGCATTTTTACCTGCAGAAAAACTTTGTATGGTACCACTGTGTTGGTTATATGGTGAA

TCTGAACGTACATCTCACTGGTATAATTATATGTAGCACTGTGCTGTGTAGATAGTTCCTAC

TGGAAAAAGAGTGGAAATTTATTAAAATCAGAAAGTATGAGATCCTGTTATGTTAAGGGAAA

TCCAAATTCCCAATTTTTTTTGGTCTTTTTAGGAAAGATTGTTGTGGTAAAAAGTGTTAGTA

TAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAAATAGTTAT

GTCTTAGGAAATTGTGGTTTAATTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTGGTTT

CATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCATCAGAATGGAACGAGTTT

TGAGTAATCAGGAAGTATATCTATATGATCTTGATATTGTTTTATAATAATTTGAAGTCTAA

AAGACTGCATTTTTAAACAAGTTAGTATTAATGCGTTGGCCCACGTAGCAAAAAGATATTTG

ATTATCTTAAAAATTGTTAAATACCGTTTTCATGAAATTTCTCAGTATTGTAACAGCAACTT

GTCAAACCTAAGCATATTTGAATATGATCTCCCATAATTTGAAATTGAAATCGTATTGTGTG

GCTCTGTATATTCTGTTAAAAAATTAAAGGACAGAAACCTTTCTTTGTGTATGCATGTTTGA

ATTAAAAGAAAGTAATGGAAG

FIGURE 15

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979

><subunit 1 of 1, 204 aa, 1 stop

><MW: 22147, pI: 8.37, NX(S/T): 3

MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVVIAVGIFLFLIALV

GLIGAVKHHQVLLFFYMIILLLVFIVQFSVSCACLALNQEQQGQLLEVGWNNTASARNDIQR

NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRFVGGIGLFFSFTEILGVWL

TYRYRNQKDPRANPSAFL

Signal Peptide:

amino acids 1-34

Transmembrane domains:

amino acids 47-63, 72-95 and 162-182

FIGURE 16

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGTNATTTTTTTATATGATTATTCTGTAANT
TGTATTTATTGTTCAGTTTTNTGTATCTTGCGCTTGTTTAGCCNTGAACCAGGAGCAACAGG
GTCAGNTTNTGGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAT
NTAAACTGCTGTGGGTTCCGAAGTGTTAACCCAAATGACACCTGTNTGGCTAGCTGTGTTAA
AAGTGACCACTNGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGA
GATTTGTTGGTGGCATTGGCCTGTTNTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACC
TACAGATACAGGAACCAG

FIGURE 17

AATCCCAAATTCCCCAATTTTTTTTGGNCTTTTTAGGGAAAGATGTGTTGTGGTAAAAAGTGT

TAGTATAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAATAG

TTATGTCTTAGGAAATTGTGGTTTAATTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTG

GTTTCATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCCATCAGAATGGAACG

AGTTTTGAGTAATCCAGGAAGTATATCTATATGATCTTGATATTGTTTTATATAATTTGAAG

TCTAAAAGACTGCATTTTTAAACAAGTTAGTATTAATGCGTTGGCCCACGTAGCAAAAAGAT

ATTTGATTATCTTAAAAATTGTTAAATACCGTTTTCATGAAAGTTCTCAGTATTGTAACAGC

AACTTGTCAAACCTAAGCATATTTGAATATGATCTCCCATAATTTGAAATTGAAATCGTATT

GTGTGGAGGAAATGGCAATCTTATGTGTGCTGAAGGACACAGTAAGAGCACCAAGTTGTGCC

CCACTTGC

FIGURE 18

ATGATTATTCTGTTACTTGTATTTATTGTTCAGTTTTATGGTATCTTGCGCTTGTTTAGCCC

CTGAAACCAGGAGCAACAGGGNNCAGCTTCCTGGAGGTTGGTTGGCAACAATCACGGCCAAG

TGACTCCGCAAATGACATCCCAGAGAAATCCTAAACTGCTGTGGGTTCCGAAGTGTTAACCC

AAATGACACCTGTCTGGCTNGCTGTGTTAAAAGTGACCACTCGTGCTCGCCATGTGCTCCAA

TCATAGGAGAATATGC

FIGURE 19

CAGTCACC**ATG**AAGCTGGGCTGTGTCCTCATGGCCTGGGCCCTCTACCTTTCCCTTGGTGTG
CTCTGGGTGGCCCAGATGCTACTGGCTGCCAGTTTTGAGACGCTGCAGTGTGAGGGACCTGT
CTGCACTGAGGAGAGCAGCTGCCACACGGAGGATGACTTGACTGATGCAAGGGAAGCTGGCT
TCCAGGTCAAGGCCTACACTTTCAGTGAACCCTTCCACCTGATTGTGTCCTATGACTGGCTG
ATCCTCCAAGGTCCAGCCAAGCCAGTTTTTGAAGGGGACCTGCTGGTTCTGCGCTGCCAGGC
CTGGCAAGACTGGCCACTGACTCAGGTGACCTTCTACCGAGATGGCTCAGCTCTGGGTCCCC
CCGGGCCTAACAGGGAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCACTACCAC
TGCAGTGGCATCTTCCAGAGCCCTGGTCCTGGGATCCCAGAAACAGCATCTGTTGTGGCTAT
CACAGTCCAAGAACTGTTTCCAGCGCCAATTCTCAGAGCTGTACCCTCAGCTGAACCCCAAG
CAGGAAGCCCCATGACCCTGAGTTGTCAGACAAAGTTGCCCCTGCAGAGGTCAGCTGCCCGC
CTCCTCTTCTCCTTCTACAAGGATGGAAGGATAGTGCAAAGCAGGGGGCTCTCCTCAGAATT
CCAGATCCCCACAGCTTCAGAAGATCACTCCGGGTCATACTGGTGTGAGGCAGCCACTGAGG
ACAACCAAGTTTGGAAACAGAGCCCCCAGCTAGAGATCAGAGTGCAGGGTGCTTCCAGCTCT
GCTGCACCTCCCACATTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAACTGCTCCTGA
GGAGGCCCCTGGGCCTCTGCCTCCGCCGCCAACCCCATCTTCTGAGGATCCAGGCTTTTCTT
CTCCTCTGGGGATGCCAGATCCTCATCTGTATCACCAGATGGGCCTTCTTCTCAAACACATG
CAGGATGTGAGAGTCCTCCTCGGTCACCTGCTCATGGAGTTGAGGGAATTATCTGGCCACCA
GAAGCCTGGGACCACAAAGGCTACTGCTGAA**TAG**AAGTAAACAGTTCATCCATGATCTCACT
TAACCACCCCAATAAATCTGATTCTTTATTTTCTCTTCCTGTCCTGCACATATGCATAAGTA
CTTTTACAAGTTGTCCCAGTGTTTTGTTAGAATAATGTAGTTAGGTGAGTGTAAATAAATTT
ATATAAAGTGAGAATTAGAGTTTAGCTATAATTGTGTATTCTCTCTTAACACAACAGAATTC
TGCTGTCTAGATCAGGAATTTCTATCTGTTATATCGACCAGAATGTTGTGATTTAAAGAGAA
CTAATGGAAGTGGATTGAATACAGCAGTCTCAACTGGGGGCAATTTTGCCCCCCAGAGGACA
TTGGGCAATGTTTGGAGACATTTTGGTCATTATACTTGGGGGGTTGGGGGATGGTGGGATGT
GTGTCTACTGGCATCCAGTAAATAGAAGCCAGGGGTGCCGCTAAACATCCTATAATGCACAG
GGCAGTACCCCACAACGAAAAATAATCTGGCCCAAAATGTCAGTTGTACTGAGTTTGAGAAA
CCCCAGCCTAATGAAACCCTAGGTGTTGGGCTCTGGAATGGGACTTTGTCCCTTCTAATTAT
TATCTCTTTCCAGCCTCATTCAGCTATTCTTACTGACATACCAGTCTTTAGCTGGTGCTATG
GTCTGTTCTTTAGTTCTAGTTTGTATCCCCTCAAAAGCCATTATGTTGAAATCCTAATCCCC
AAGGTGATGGCATTAAGAAGTGGGCCTTTGGGAAGTGATTAGATCAGGAGTGCAGAGCCCTC
ATGATTAGGATTAGTGCCCTTATTTAAAAAGGCCCCAGAGAGCTAACTCACCCTTCCACCAT
ATGAGGACGTGGCAAGAAGATGACATGTATGAGAACCAAAAAACAGCTGTCGCCAAACACCG
ACTCTGTCGTTGCCTTGATCTTGAACTTCCAGCCTCCAGAACTATGAGAAATAAAATTCTGG
TTGTTTGTAGCCTAA

FIGURE 20

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594

><subunit 1 of 1, 359 aa, 1 stop

><MW: 38899, pI: 5.21, NX(S/T): 0

MKLGCVLMAWALYLSLGVLWVAQMLLAASFETLQCEGPVCTEESSCHTEDDLTDAREAGFQV

KAYTFSEPFHLIVSYDWLILQGPAKPVFEGDLLVLRCQAWQDWPLTQVTFYRDGSALGPPGP

NREFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGS

PMTLSCQTKLPLQRSAARLLFSFYKDGRIVQSRGLSSEFQIPTASEDHSGSYWCEAATEDNQ

VWKQSPQLEIRVQGASSSAAPPTLNPAPQKSAAPGTAPEEAPGPLPPPPTPSSEDPGFSSPL

GMPDPHLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE

Signal sequence:

amino acids 1-17

Leucine zipper pattern sequence:

amino acids 12-33

Protein kinase C phosphorylation site:

amino acids 353-355

FIGURE 21

CCCACGCGTCCGCCCACGCGTCCGCCCACGGGTCCGCCCACGCGTCCGGGCCACCAGAAGTT
TGAGCCTCTTTGGTAGCAGGAGGCTGGAAGAAAGGACAGAAGTAGCTCTGGCTGTG**ATG**GGG
ATCTTACTGGGCCTGCTACTCCTGGGGCACCTAACAGTGGACACTTATGGCCGTCCCATCCT
GGAAGTGCCAGAGAGTGTAACAGGACCTTGGAAAGGGGATGTGAATCTTCCCTGCACCTATG
ACCCCCTGCAAGGCTACACCCAAGTCTTGGTGAAGTGGCTGGTACAACGTGGCTCAGACCCT
GTCACCATCTTTCTACGTGACTCTTCTGGAGACCATATCCAGCAGGCAAAGTACCAGGGCCG
CCTGCATGTGAGCCACAAGGTTCCAGGAGATGTATCCCTCCAATTGAGCACCCTGGAGATGG
ATGACCGGAGCCACTACACGTGTGAAGTCACCTGGCAGACTCCTGATGGCAACCAAGTCGTG
AGAGATAAGATTACTGAGCTCCGTGTCCAGAAACTCTCTGTCTCCAAGCCCACAGTGACAAC
TGGCAGCGGTTATGGCTTCACGGTGCCCCAGGGAATGAGGATTAGCCTTCAATGCCAGGCTC
GGGGTTCTCCTCCCATCAGTTATATTTGGTATAAGCAACAGACTAATAACCAGGAACCCATC
AAAGTAGCAACCCTAAGTACCTTACTCTTCAAGCCTGCGGTGATAGCCGACTCAGGCTCCTA
TTTCTGCACTGCCAAGGGCCAGGTTGGCTCTGAGCAGCACAGCGACATTGTGAAGTTTGTGG
TCAAAGACTCCTCAAAGCTACTCAAGACCAAGACTGAGGCACCTACAACCATGACATACCCC
TTGAAAGCAACATCTACAGTGAAGCAGTCCTGGGACTGGACCACTGACATGGATGGCTACCT
TGGAGAGACCAGTGCTGGGCCAGGAAAGAGCCTGCCTGTCTTTGCCATCATCCTCATCATCT
CCTTGTGCTGTATGGTGGTTTTTACCATGGCCTATATCATGCTCTGTCGGAAGACATCCCAA
CAAGAGCATGTCTACGAAGCAGCCAGG**TAA**GAAAGTCTCTCCTCTTCCATTTTTGACCCCGT
CCCTGCCCTCAATTTTGATTACTGGCAGGAAATGTGGAGGAAGGGGGGTGTGGCACAGACCC
AATCCTAAGGCCGGAGGCCTTCAGGGTCAGGACATAGCTGCCTTCCCTCTCTCAGGCACCTT
CTGAGGTTGTTTTGGCCCTCTGAACACAAAGGATAATTTAGATCCATCTGCCTTCTGCTTCC
AGAATCCCTGGGTGGTAGGATCCTGATAATTAATTGGCAAGAATTGAGGCAGAAGGGTGGGA
AACCAGGACCACAGCCCCAAGTCCCTTCTTATGGGTGGTGGGCTCTTGGGCCATAGGGCACA
TGCCAGAGAGGCCAACGACTCTGGAGAAACCATGAGGGTGGCCATCTTCGCAAGTGGCTGCT
CCAGTGATGAGCCAACTTCCCAGAATCTGGGCAACAACTACTCTGATGAGCCCTGCATAGGA
CAGGAGTACCAGATCATCGCCCAGATCAATGGCAACTACGCCCGCCTGCTGGACACAGTTCC
TCTGGATTATGAGTTTCTGGCCACTGAGGGCAAAAGTGTCTGTTAAAAATGCCCCATTAGGC
CAGGATCTGCTGACATAATTGCCTAGTCAGTCCTTGCCTTCTGCATGGCCTTCTTCCCTGCT
ACCTCTCTTCCTGGATAGCCCAAAGTGTCCGCCTACCAACACTGGAGCCGCTGGGAGTCACT
GGCTTTGCCCTGGAATTTGCCAGATGCATCTCAAGTAAGCCAGCTGCTGGATTTGGCTCTGG
GCCCTTCTAGTATCTCTGCCGGGGGCTTCTGGTACTCCTCTCTAAATACCAGAGGGAAGATG
CCCATAGCACTAGGACTTGGTCATCATGCCTACAGACACTATTCAACTTTGGCATCTTGCCA
CCAGAAGACCCGAGGGAGGCTCAGCTCTGCCAGCTCAGAGGACCAGCTATATCCAGGATCAT
TTCTCTTTCTTCAGGGCCAGACAGCTTTTAATTGAAATTGTTATTTCACAGGCCAGGGTTCA
GTTCTGCTCCTCCACTATAAGTCTAATGTTCTGACTCTCTCCTGGTGCTCAATAAATATCTA
ATCATAACAGC

FIGURE 22

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45416

><subunit 1 of 1, 321 aa, 1 stop

><MW: 35544, pI: 8.51, NX(S/T): 0

MGILLGLLLLGHLTVDTYGRPILEVPESVTGPWKGDVNLPCTYDPLQGYTQVLVKWLVQRGS

DPVTIFLRDSSGDHIQQAKYQGRLHVSHKVPGDVSLQLSTLEMDDRSHYTCEVTWQTPDGNQ

VVRDKITELRVQKLSVSKPTVTTGSGYGFTVPQGMRISLQCQARGSPPISYIWYKQQTNNQE

PIKVATLSTLLFKPAVIADSGSYFCTAKGQVGSEQHSDIVKFVVKDSSKLLKTKTEAPTTMT

YPLKATSTVKQSWDWTTDMDGYLGETSAGPGKSLPVFAIILIISLCCMVVFTMAYIMLCRKT

SQQEHVYEAAR

Signal Sequence:

amino acids 1-19

Glycosaminoglycan attachment site:

amino acids 149-152

Transmembrane domain:

amino acids 282-300

FIGURE 23

GCGCCGGGAGCCCATCTGCCCCCAGGGGCACGGGGCGCGGGGCCGGCTCCCGCCCGGCACAT
GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCGCGCCCAGCTCGCCCGAGGTCCGTCGGA
GGCGCCCGGCCGCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATC
GGG**ATG**TCCCTCCTCCTTCTCCTCTTGCTAGTTTCCTACTATGTTGGAACCTTGGGGACTCA
CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGC
TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA
GTGGTGATCACTTACTCCAGTCGTCATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG
AGTGGCCTTTGCTTCCAATTTCCTGGCAGGAGATGCCTCCTTGCAGATTGAACCTCTGAAGC
CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTCAGGGCGCTACGTGTGGAGCCAT
GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC
AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATT
ACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT
GACTACAACCACCCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA
CCAGTGCACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAACTGTACAGT
ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG
ATTTTCCTCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA
GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT
CCTCTTCCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCCTCCACTCGCTCCACAGCAAAT
AGTGCCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCCAGCCAGGGCTGGCCAC
CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
CTAATCTGACCAAAGCAGAAACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAA
ACGGTC**TGA**ATTACAATGGACTTGACTCCCACGCTTTCCTAGGAGTCAGGGTCTTTGGACTC
TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA
GTGAGCATTGCACGGAACAGATTCAGATGAGCATTTTCCTTATACAATACCAAACAAGCAAA
AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG
AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG
AGGTGAATATACCTAAAACTTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTCACAATT
TTCAAGAGGAAATGGGATGCTGTTTGTAAATTTTCTATGCATTTCTGCAAACTTATTGGATT
ATTAGTTATTCAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
TGAGCTAACCACTTCTAAGAAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC
TTCATTTGTCATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA
AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAAATAAC
TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCCATCTTCATGATGTT
ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCCCTCAAAT
CAGATGCCTCTAAGGACTTTCCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT
TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA
CCCAACATACCATTATAGTCTCTTCTTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG
GTGGACTAGAAAGGGAGATTAGATCAGTTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA
TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV

VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV

ILKVLVRPSKPKCELEGELTEGSDLTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID

YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI

FLLVWLLIRRKDKERYEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS

ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

FIGURE 25

GTCGTTCCTTTGCTCTCTCGCGCCCAGTCCTCCTCCCTGGTTCTCCTCAGCCGCTGTCGGAGGAGAGCACCCGGA
GACGCGGGCTGCAGTCGCGGCGGCTTCTCCCCGCCTGGGCGGCCTCGCCGCTGGGCAGGTGCTGAGCGCCCCTAG
AGCCTCCCTTGCCGCCTCCCTCCTCTGCCCGGCCGCAGCAGTGCACATGGGGTGTTGGAGGTAGATGGGCTCCCG
GCCCGGGAGGCGGCGGTGGATGCGGCGCTGGGCAGAAGCAGCCGCCGATTCCAGCTGCCCCGCGCGCCCCGGGCG
CCCCTGCGAGTCCCCGGTTCAGCC**ATG**GGGACCTCTCCGAGCAGCAGCACCGCCCTCGCCTCCTGCAGCCGCATC
GCCCGCCGAGCCACAGCCACGATGATCGCGGGCTCCCTTCTCCTGCTTGGATTCCTTAGCACCACCACAGCTCAG
CCAGAACAGAAGGCCTCGAATCTCATTGGCACATACCGCCATGTTGACCGTGCCACCGGCCAGGTGCTAACCTGT
GACAAGTGTCCAGCAGGAACCTATGTCTCTGAGCATTGTACCAACACAAGCCTGCGCGTCTGCAGCAGTTGCCCT
GTGGGGACCTTTACCAGGCATGAGAATGGCATAGAGAAATGCCATGACTGTAGTCAGCCATGCCCATGGCCAATG
ATTGAGAAATTACCTTGTGCTGCCTTGACTGACCGAGAATGCACTTGCCCACCTGGCATGTTCCAGTCTAACGCT
ACCTGTGCCCCCCATACGGTGTGTCCTGTGGGTTGGGGTGTGCGGAAGAAAGGGACAGAGACTGAGGATGTGCGG
TGTAAGCAGTGTGCTCGGGGTACCTTCTCAGATGTGCCTTCTAGTGTGATGAAATGCAAAGCATACACAGACTGT
CTGAGTCAGAACCTGGTGGTGATCAAGCCGGGGACCAAGGAGACAGACAACGTCTGTGGCACACTCCCGTCCTTC
TCCAGCTCCACCTCACCTTCCCCTGGCACAGCCATCTTTCCACGCCCTGAGCACATGGAAACCCATGAAGTCCCT
TCCTCCACTTATGTTCCCAAAGGCATGAACTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACTG
AGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGGGGAAGGAAGACGTGAACAAGACCCTC
CCAAACCTTCAGGTAGTCAACCACCAGCAAGGCCCCCACCACAGACACATCCTGAAGCTGCTGCCGTCCATGGAG
GCCACTGGGGGCGAGAAGTCCAGCACGCCCATCAAGGGCCCCAAGAGGGGACATCCTAGACAGAACCTACACAAG
CATTTTGACATCAATGAGCATTTGCCCTGGATGATTGTGCTTTTCCTGCTGCTGGTGCTTGTGGTGATTGTGGTG
TGCAGTATCCGGAAAAGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGGATCCCAGTGCCATTGTGGAAAAGGCA
GGGCTGAAGAAATCCATGACTCCAACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGAT
ATCCTGAAGCTTGTAGCAGCCCAAGTGGGAAGCCAGTGGAAAGATATCTATCAGTTTCTTTGCAATGCCAGTGAG
AGGGAGGTTGCTGCTTTCTCCAATGGGTACACAGCCGACCACGAGCGGGCCTACGCAGCTCTGCAGCACTGGACC
ATCCGGGGCCCCGAGGCCAGCCTCGCCCAGCTAATTAGCGCCCTGCGCCAGCACCGGAGAAACGATGTTGTGGAG
AAGATTCGTGGGCTGATGGAAGACACCACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCCG
CTTAGCCCCGAGCCCCATCCCCCAGCCCAACGCGAAACTTGAGAATTCCGCTCTCCTGACGGTGGAGCCTTCCCCA
CAGGACAAGAACAAGGGCTTCTTCGTGGATGAGTCGGAGCCCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTCC
TCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAAGGACACAGTGTTGCGGCAGGTACGCCTGGAC
CCCTGTGACTTGCAGCCTATCTTTGATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGAG
ATTCCCCAGGCTGAGGACAAACTAGACCGGCTATTCGAAATTATTGGAGTCAAGAGCCAGGAAGCCAGCCAGACC
CTCCTGGACTCTGTTTATAGCCATCTTCCTGACCTGCTG**TAG**AACATAGGGATACTGCATTCTGGAAATTACTCA
ATTTAGTGGCAGGGTGGTTTTTTAATTTTCTTCTGTTTCTGATTTTTGTTGTTTGGGGTGTGTGTGTGTGTTTGT
GTGTGTGTGTGTGTGTGTGTGTGTGTGTTTAACAGAGAATATGGCCAGTGCTTGAGTTCTTTCTCCTTCTC
TCTCTCTCTTTTTTTTTAAATAACTCTTCTGGGAAGTTGGTTTATAAGCCTTTGCCAGGTGTAACTGTTGTGAA
ATACCCACCACTAAAGTTTTTTAAGTTCCATATTTTCTCCATTTTGCCTTCTTATGTATTTTCAAGATTATTCTG
TGCACTTTAAATTTACTTAACTTACCATAAATGCAGTGTGACTTTTCCCACACACTGGATTGTGAGGCTCTTAAC
TTCTTAAAAGTATAATGGCATCTTGTGAATCCTATAAGCAGTCTTTATGTCTCTTAACATTCACACCTACTTTTT
AAAAACAAATATTATTACTATTTTTATTATTGTTTGTCCTTTATAAATTTTCTTAAAGATTAAGAAAATTTAAGA
CCCCATTGAGTTACTGTAATGCAATTCAACTTTGAGTTATCTTTTAAATATGTCTTGTATAGTTCATATTCATGG
CTGAAACTTGACCACACTATTGCTGATTGTATGGTTTTCACCTGGACACCGTGTAGAATGCTTGATTACTTGTAC
TCTTCTTATGCTAATATGCTCTGGGCTGGAGAAATGAAATCCTCAAGCCATCAGGATTTGCTATTTAAGTGGCTT
GACAACTGGGCCACCAAAGAACTTGAACTTCACCTTTTAGGATTTGAGCTGTTCTGGAACACATTGCTGCACTTT
GGAAAGTCAAAATCAAGTGCCAGTGGCGCCCTTTCCATAGAGAATTTGCCCAGCTTTGCTTTAAAGATGTCTTG
TTTTTTATATACACATAATCAATAGGTCCAATCTGCTCTCAAGGCCTTGGTCCTGGTGGGATTCCTTCACCAATT
ACTTTAATTAAAAATGGCTGCAACTGTAAGAACCCTTGTCTGATATATTTGCAACTATGCTCCCATTTACAAATG
TACCTTCTAATGCTCAGTTGCCAGGTTCCAATGCAAAGGTGGCGTGGACTCCCTTTGTGTGGGTGGGGTTTGTGG
GTAGTGGTGAAGGACCGATATCAGAAAAATGCCTTCAAGTGTACTAATTTATTAATAAACATTAGGTGTTTGTTA
AAAAAAAAA

FIGURE 26

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52594

><subunit 1 of 1, 655 aa, 1 stop

><MW: 71845, pI: 8.22, NX(S/T): 8

MGTSPSSSTALASCSRIARRATATMIAGSLLLLGFLSTTTAQPEQKASNLIGTYRHVDRATG

QVLTCDKCPAGTYVSEHCTNTSLRVCSSCPVGTFTRHENGIEKCHDCSQPCPWPMIEKLPCA

ALTDRECTCPPGMFQSNATCAPHTVCPVGWGVRKKGTETEDVRCKQCARGTFSDVPSSVMKC

KAYTDCLSQNLVVIKPGTKETDNVCGTLPSFSSSTSPSPGTAIFPRPEHMETHEVPSSTYVP

KGMNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL

KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSIRK

SSRTLKKGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAAQVGSQWKDIY

QFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDVVEKIRG

LMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNKGFFVDESEPLL

RCDSTSSGSSALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPEELRVIEEIPQ

AEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 350-370

FIGURE 27

ATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCCGTGGTGCCATCTACATTTTTGGGA
CTCGGGAATTATGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAAATAGTCAC
C**ATG**GGGGAAAATGATCCGCCTGCTGTTGAAGCCCCCTTCTCATTCCGATCGCTTTTTGGCC
TTGATGATTTGAAAATAAGTCCTGTTGCACCAGATGCAGATGCTGTTGCTGCACAGATCCTG
TCACTGCTGCCATTGAAGTTTTTTCCAATCATCGTCATTGGGATCATTGCATTGATATTAGC
ACTGGCCATTGGTCTGGGCATCCACTTCGACTGCTCAGGGAAGTACAGATGTCGCTCATCCT
TTAAGTGTATCGAGCTGATAGCTCGATGTGACGGAGTCTCGGATTGCAAAGACGGGGAGGAC
GAGTACCGCTGTGTCCGGGTGGGTGGTCAGAATGCCGTGCTCCAGGTGTTCACAGCTGCTTC
GTGGAAGACCATGTGCTCCGATGACTGGAAGGGTCACTACGCAAATGTTGCCTGTGCCCAAC
TGGGTTTCCCAAGCTATGTGAGTTCAGATAACCTCAGAGTGAGCTCGCTGGAGGGGCAGTTC
CGGGAGGAGTTTGTGTCCATCGATCACCTCTTGCCAGATGACAAGGTGACTGCATTACACCA
CTCAGTATATGTGAGGGAGGGATGTGCCTCTGGCCACGTGGTTACCTTGCAGTGCACAGCCT
GTGGTCATAGAAGGGGCTACAGCTCACGCATCGTGGGTGGAAACATGTCCTTGCTCTCGCAG
TGGCCCTGGCAGGCCAGCCTTCAGTTCCAGGGCTACCACCTGTGCGGGGGCTCTGTCATCAC
GCCCCTGTGGATCATCACTGCTGCACACTGTGTTTATGACTTGTACCTCCCCAAGTCATGGA
CCATCCAGGTGGGTCTAGTTTCCCTGTTGGACAATCCAGCCCCATCCCACTTGGTGGAGAAG
ATTGTCTACCACAGCAAGTACAAGCCAAAGAGGCTGGGCAATGACATCGCCCTTATGAAGCT
GGCCGGGCCACTCACGTTCAATGAAATGATCCAGCCTGTGTGCCTGCCCAACTCTGAAGAGA
ACTTCCCCGATGGAAAAGTGTGCTGGACGTCAGGATGGGGGGGCCACAGAGGATGGAGGTGAC
GCCTCCCCTGTCCTGAACCACGCGGCCGTCCCTTTGATTTCCAACAAGATCTGCAACCACAG
GGACGTGTACGGTGGCATCATCTCCCCCTCCATGCTCTGCGCGGGCTACCTGACGGGTGGCG
TGGACAGCTGCCAGGGGGACAGCGGGGGGCCCCTGGTGTGTCAAGAGAGGAGGCTGTGGAAG
TTAGTGGGAGCGACCAGCTTTGGCATCGGCTGCGCAGAGGTGAACAAGCCTGGGGTGTACAC
CCGTGTCACCTCCTTCCTGGACTGGATCCACGAGCAGATGGAGAGAGACCTAAAAACC**TGAA**
GAGGAAGGGGACAAGTAGCCACCTGAGTTCCTGAGGTGATGAAGACAGCCCGATCCTCCCCT
GGACTCCCGTGTAGGAACCTGCACACGAGCAGACACCCTTGGAGCTCTGAGTTCCGGCACCA
GTAGCAGGCCCGAAAGAGGCACCCTTCCATCTGATTCCAGCACAACCTTCAAGCTGCTTTTT
GTTTTTTGTTTTTTTGAGGTGGAGTCTCGCTCTGTTGCCCAGGCTGGAGTGCAGTGGCGAAA
TCCCTGCTCACTGCAGCCTCCGCTTCCCTGGTTCAAGCGATTCTCTTGCCTCAGCTTCCCCA
GTAGCTGGGACCACAGGTGCCCGCCACCACACCCAACTAATTTTTGTATTTTTAGTAGAGAC
AGGGTTTCACCATGTTGGCCAGGCTGCTCTCAAACCCCTGACCTCAAATGATGTGCCTGCTT
CAGCCTCCCACAGTGCTGGGATTACAGGCATGGGCCACCACGCCTAGCCTCACGCTCCTTTC
TGATCTTCACTAAGAACAAAAGAAGCAGCAACTTGCAAGGGCGGCCTTTCCCACTGGTCCAT
CTGGTTTTCTCTCCAGGGTCTTGCAAAATTCCTGACGAGATAAGCAGTTATGTGACCTCACG
TGCAAAGCCACCAACAGCCACTCAGAAAAGACGCACCAGCCCAGAAGTGCAGAACTGCAGTC
ACTGCACGTTTTCATCTCTAGGGACCAGAACCAAACCCACCCTTTCTACTTCCAAGACTTAT
TTTCACATGTGGGGAGGTTAATCTAGGAATGACTCGTTTAAGGCCTATTTTCATGATTTCTT
TGTAGCATTTGGTGCTTGACGTATTATTGTCCTTTGATTCCAAATAATATGTTTCCTTCCCT
CATTGTCTGGCGTGTCTGCGTGGACTGGTGACGTGAATCAAAATCATCCACTGAAA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234

><subunit 1 of 1, 453 aa, 1 stop

><MW: 49334, pI: 6.32, NX(S/T): 1

MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILA
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVRVGGQNAVLQVFTAAS
WKTMCSDDWKGHYANVACAQLGFPSYVSSDNLRVSSLEGQFREEFVSIDHLLPDDKVTALHH
SVYVREGCASGHVVTLQCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVIT
PLWIITAAHCVYDLYLPKSWTIQVGLVSLLDNPAPSHLVEKIVYHSKYKPKRLGNDIALMKL
AGPLTFNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAVPLISNKICNHR
DVYGGIISPSMLCAGYLTGGVDSCQGDSGGPLVCQERRLWKLVGATSFGIGCAEVNKPGVYT
RVTSFLDWIHEQMERDLKT

Signal Peptide:

amino acids 1-20

Transmembrane domain:

amino acids 240-284

FIGURE 29

CCCACGCGTCCGTCCTAGTCCCCGGGCCAACTCGGACAGTTTGCTCATTTATTGCAACGGTCAAGGCTGGCTTGT
GCCAGAACGGCGCGCGCGCGCGCACGCACGCACACACACGGGGGGAAACTTTTTTAAAAATGAAAGGCTAGAAGA
GCTCAGCGGCGGCGCGGGCGCTGCGCGAGGGCTCCGGAGCTGACTCGCCGAGGCAGGAAATCCCTCCGGTCGCGA
CGCCCGGCCCCGGCTCGGCGCCCGCGTGGGATGGTGCAGCGCTCGCCGCCGGGCCCGAGAGCTGCTGCACTGAAG
GCCGGCGACG<u>**ATG**</u>GCAGCGCGCCCGCTGCCCGTGTCCCCCGCCCGCGCCCTCCTGCTCGCCCTGGCCGGTGCTCT
GCTCGCGCCCTGCGAGGCCCGAGGGGTGAGCTTATGGAACCAAGGAAGAGCTGATGAAGTTGTCAGTGCCTCTGT
TCGGAGTGGGGACCTCTGGATCCCAGTGAAGAGCTTCGACTCCAAGAATCATCCAGAAGTGCTGAATATTCGACT
ACAACGGGAAAGCAAAGAACTGATCATAAATCTGGAAAGAAATGAAGGTCTCATTGCCAGCAGTTTCACGGAAAC
CCACTATCTGCAAGACGGTACTGATGTCTCCCTCGCTCGAAATTACACGGGTCACTGTTACTACCATGGACATGT
ACGGGGATATTCTGATTCAGCAGTCAGTCTCAGCACGTGTTCTGGTCTCAGGGGACTTATTGTGTTTGAAAATGA
AAGCTATGTCTTAGAACCAATGAAAAGTGCAACCAACAGATACAAACTCTTCCCAGCGAAGAAGCTGAAAAGCGT
CCGGGGGATCATGTGGATCACATCACAACACACCAAACCTCGCTGCAAAGAATGTGTTTCCACCACCCTCTCAGAC
ATGGGCAAGAAGGCATAAAGAGAGACCCTCAAGGCAACTAAGTATGTGGAGCTGGTGATCGTGGCAGACAACCG
AGAGTTTCAGAGGCAAGGAAAAGATCTGGAAAAAGTTAAGCAGCGATTAATAGAGATTGCTAATCACGTTGACAA
GTTTTACAGACCACTGAACATTCGGATCGTGTTGGTAGGCGTGGAAGTGTGGAATGACATGGACAAATGCTCTGT
AAGTCAGGACCCATTCACCAGCCTCCATGAATTTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAAATCCCA
TGACAATGCGCAGCTTGTCAGTGGGGTTTATTTCCAAGGGACCACCATCGGCATGGCCCCAATCATGAGCATGTG
CACGGCAGACCAGTCTGGGGGAATTGTCATGGACCATTCAGACAATCCCCTTGGTGCAGCCGTGACCCTGGCACA
TGAGCTGGGCCACAATTTCGGGATGAATCATGACACACTGGACAGGGGCTGTAGCTGTCAAATGGCGGTTGAGAA
AGGAGGCTGCATCATGAACGCTTCCACCGGGTACCCATTTCCCATGGTGTTCAGCAGTTGCAGCAGGAAGGACTT
GGAGACCAGCCTGGAGAAAGGAATGGGGGTGTGCCTGTTTAACCTGCCGGAAGTCAGGGAGTCTTTCGGGGGCCA
GAAGTGTGGGAACAGATTTGTGGAAGAAGGAGAGGAGTGTGACTGTGGGGAGCCAGAGGAATGTATGAATCGCTG
CTGCAATGCCACCACCTGTACCCTGAAGCCGGACGCTGTGTGCGCACATGGGCTGTGCTGTGAAGACTGCCAGCT
GAAGCCTGCAGGAACAGCGTGCAGGGACTCCAGCAACTCCTGTGACCTCCCAGAGTTCTGCACAGGGGCCAGCCC
TCACTGCCCAGCCAATGTGTACCTGCACGATGGGCACTCATGTCAGGATGTGGACGGCTACTGCTACAATGGCAT
CTGCCAGACTCACGAGCAGCAGTGTGTCACGCTCTGGGGACCAGGTGCTAAACCTGCCCCTGGGATCTGCTTTGA
GAGAGTCAATTCTGCAGGTGATCCTTATGGCAACTGTGGCAAAGTCTCGAAGAGTTCCTTTGCCAAATGCGAGAT
GAGAGATGCTAAATGTGGAAAAAATCCAGTGTCAAGGAGGTGCCAGCCGGCCAGTCATTGGTACCAATGCCGTTTC
CATAGAAACAAACATCCCTCTGCAGCAAGGAGGCCGGATTCTGTCCGGGGGACCCCACGTGTACTTGGGCGATGA
CATGCCGGACCCAGGGCTTGTGCTTGCAGGCACAAAGTGTGCAGATGGAAAAATCTGCCTGAATCGTCAATGTCA
AAATATTAGTGTCTTTGGGGTTCACGAGTGTGCAATGCAGTGCCACGGCAGAGGGGTGTGCAACAACAGGAAGAA
CTGCCACTGCGAGGCCCACTGGGCACCTCCCTTCTGTGACAAGTTTGGCTTTGGAGGAAGCACAGACAGCGGCCC
CATCCGGCAAGCAGAAGCAAGGCAGGAAGCTGCAGAGTCCAACAGGGAGCGCGGCCAGGGCCAGGAGCCCGTGGG
ATCGCAGGAGCATGCGTCTACTGCCTCACTGACACTCATC<u>**TGA**</u>GCCCTCCCATGACATGGAGACCGTGACCAGTG
CTGCTGCAGAGGAGGTCACGCGTCCCCAAGGCCTCCTGTGACTGGCAGCATTGACTCTGTGGCTTTGCCATCGTT
TCCATGACAACAGACACAACACAGTTCTCGGGGCTCAGGAGGGGAAGTCCAGCCTACCAGGCACGTCTGCAGAAA
CAGTGCAAGGAAGGGCAGCGACTTCCTGGTTGAGCTTCTGCTAAAACATGGACATGCTTCAGTGCTGCTCCTGAG
AGAGTAGCAGGTTACCACTCTGGCAGGCCCCAGCCCTGCAGCAAGGAGGAAGAGGACTCAAAAGTCTGGCCTTTC
ACTGAGCCTCCACAGCAGTGGGGGAGAAGCAAGGGTTGGGCCCAGTGTCCCCTTTCCCCAGTGACACCTCAGCCT
TGGCAGCCCTGATGACTGGTCTCTGGCTGCAACTTAATGCTCTGATATGGCTTTTAGCATTTATTATATGAAAAT
AGCAGGGTTTTAGTTTTTAATTTATCAGAGACCCTGCCACCCATTCCATCTCCATCCAAGCAAACTGAATGGCAA
TGAAACAAACTGGAGAAGAAGGTAGGAGAAAGGGCGGTGAACTCTGGCTCTTTGCTGTGGACATGCGTGACCAGC
AGTACTCAGGTTTGAGGGTTTGCAGAAAGCCAGGGAACCCACAGAGTCACCAACCCTTCATTTAACAAGTAAGAA
TGTTAAAAAGTGAAAACAATGTAAGAGCCTAACTCCATCCCCGTGGCCATTACTGCATAAAATAGAGTGCATTT
GAAAT

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624

><subunit 1 of 1, 735 aa, 1 stop

><MW: 80177, pI: 7.08, NX(S/T): 5

MAARPLPVSPARALLLALAGALLAPCEARGVSLWNQGRADEVVSASVRSGDLWIPVKSFDSK

NHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTGHCYYHGHVRG

YSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSVRGSCGSHHNTPNLA

AKNVFPPPSQTWARRHKRETLKATKYVELVIVADNREFQRQGKDLEKVKQRLIEIANHVDKF

YRPLNIRIVLVGVEVWNDMDKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQLVSGVYFQG

TTIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNFGMNHDTLDRGCSCQMAVEK

GGCIMNASTGYPFPMVFSSCSRKDLETSLEKGMGVCLFNLPEVRESFGGQKCGNRFVEEGEE

CDCGEPEECMNRCCNATTCTLKPDAVCAHGLCCEDCQLKPAGTACRDSSNSCDLPEFCTGAS

PHCPANVYLHDGHSCQDVDGYCYNGICQTHEQQCVTLWGPGAKPAPGICFERVNSAGDPYGN

CGKVSKSSFAKCEMRDAKCGKIQCQGGASRPVIGTNAVSIETNIPLQQGGRILCRGTHVYLG

DDMPDPGLVLAGTKCADGKICLNRQCQNISVFGVHECAMQCHGRGVCNNRKNCHCEAHWAPP

FCDKFGFGGSTDSGPIRQAEARQEAAESNRERGQGQEPVGSQEHASTASLTLI

Signal peptide:

amino acids 1-28

FIGURE 31

TCCCAAGGCTTCTTGGATGGCAGATGATTNTGGGGTTTTGCATTGTTTCCCTGACAACGAAA

ACAAAACAGTTTTGGGGGTTCAGGAGGGGAANTCCAGCCTACCCAGGAAGTTTGCAGAAACA

GTGCAAGGAAGGGCAGGANTTCCTGGTTGAGNTTTTTGNTAAAACATGGACATGNTTCAGTG

CTGCTCNTGAGAGAGTAGCAGGTTACCACTTTTGGCAGGCCCCAGCCCTGCAGCAAGGAGGA

AGAGGACTCAAAAGTTTGGCCTTTCACTGAGCCTCCACAGCAGTGGGGGAGAAGCAAGGGTT

GGGCCCAGTGTCCCCTTTCCCCAGTGACACCTCAGCCTTGGCAGCCCTGATAACTGGTNTNT

GGCTGCAANTTAATGCTNTGATATGGCTTTTAGCATTTATTATATGAAAATAGCAGGGTTTT

AGTTTTTAATTTATCAGAGACCCTGCCACCCATTCCATNTCCATCCAAG

FIGURE 32

CATCCTGCAACATGGTGAAACCACGCCTGGCTAATTTTGTTGTATTTTTGGTAGAGATGGGA

TTTCACCGTGTTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCGCCTCGGCCTCCC

AAAGTGCTGGGATTACAGGCGAGTGCAACCACACCCGGCCACAAACTTTTTAAGAAGTTAAT

GAAACCATACCTTTTACATTTTTAATGACAGGAAAATGCTCACAATAATTGTTAACCCAAAA

TTCTGGATACAAAAGTACAATCTTTACTGTGTAAATACATGTATATGTACTATATGAAAATA

TACCAAATATCAATAATACTTATCTCTGGGTAAAAACCTCTTCTCATACCCTGTGCTAACAA

CTTTTAACAAAAATTTGCATCACTTTTAAGAATCAAGAAAAATTTCTGAAGGTCATATGGG

ACAGAAAAAAAACCAAGGGAAAAATCACGCCACTTGGGAAAAAAAGATTCGAAATCTGCCT

TTTTATAGATTTGTAATTAATAAGGTCCAGGCTTTCTAAGCAACTTAAATGTTTTGTTTCGA

AACAAAGTACTTGTCTGGATGTAGGAGGAAAGGGAGTGATGTCACTGCCATTATGATGCCCC

TTGAATATAAGACCCTACTTGCTATCTCCCCTGCACCAGCCAGGAGCCACCCATCCTCCAGC

ACACTGAGCAGCAAGCTGGACACACGGCACACTGATCCAA**ATG**GGTAAGGGGATGGTGGCGA

TGCTCATTCTGGGTCTGCTACTTCTGGCGCTGCTCCTACCCGTGCAGGTTTCTTCATTTGTT

CCTTTAACCAGTATGCCGGAAGCTACTGCAGCCGAAACCACAAAGCCCTCCAACAGTGCCCT

ACAGCCTACAGCCGGTCTCCTTGTGGTCTTGCTTGCCCTTCTACATCTCTACCAT**TAA**GAGG

CAGGTCAAGAAACAGCTACAGTTCTCCAACCCATACACTAAAACCGAATCCAAATGGTGCCT

AGAAGTTCAATGTGGCAAGGAAAAAAACCAGGTCTTCATCAAATCTACTAATTTCACTCCTT

ATTAACAGAGAAACGCTTGAGAGTCTCAAACTGGACTGGTTTAAAGAGCATCTGAAGGATTT

GACTAGATGATAAATGCCTGTACTCCCAGTACTTTGGGAGGCCTAGGCCGGCGGATCACCTG

AGGTCAGGAGTTTGAGACTAACCTGGCCAAAATGGTGAAACCCCATCTGTACTAAAAATACA

AATATTGACTGGGCGTGGTGGTGAGTGCCTGTGATCCCAGCTACTCAGGTGGCTGAAGCAGG

ACAATCACTTGAACTCAGGAGGCAGAGGTTGCAGTGAGCTGAGATCGCGCTACTGCACTCTA

GCCTAGCCTGGGCAACAGAGTGAGACTTCGTCTCAAAAAAAAAAAAAGCCAAGTGCAGTGGCT

CACGCCTGTAATCCCGGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAGGTCAGGAGATCA

AGACCATCCTGGCTAATACAGTGAAACCCTGTCTCTACTAAAAATACAAAAAATTAGCCGGG

GATGGTGGCAGGCACCTGGAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATAGCGTGAA

CTCAGGAGGCGGAGCTTGCAGTGAGCCGAGATTGCGCTACTGCACTCCAGCCTGGGCGACAG

CGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 33

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48309

><subunit 1 of 1, 67 aa, 1 stop

><MW: 6981, pI: 7.47, NX(S/T): 0

MGKGMVAMLILGLLLLALLLPVQVSSFVPLTSMPEATAAETTKPSNSALQPTAGLLVVLLAL
LHLYH

Signal peptide:

amino acids 15-27

FIGURE 34

GCCGCGGCGAGAGCGCGCCCAGCCCCGCCGCG**ATG**CCCGCGCGCCCAGGACGCCTCCTCCCGCTGCTGGCCCGGC
CGGCGGCCCTGACTGCGCTGCTGCTGCTGCTGCTGGGCCATGGCGGCGGCGGGCGCTGGGGCGCCCGGGCCCAGG
AGGCGGCGGCGGCGGCGGACGGGCCCCCCGCGGCAGACGGCGAGGACGGACAGGACCCGCACAGCAAGCACC
TGTACACGGCCGACATGTTCACGCACGGGATCCAGAGCGCCGCGCACTTCGTCATGTTCTTCGCGCCCTGGTGTG
GACACTGCCAGCGGCTGCAGCCGACTTGGAATGACCTGGGAGACAAATACAACAGCATGGAAGATGCCAAAGTCT
ATGTGGCTAAAGTGGACTGCACGGCCCACTCCGACGTGTGCTCCGCCCAGGGGGTGCGAGGATACCCCACCTTAA
AGCTTTTCAAGCCAGGCCAAGAAGCTGTGAAGTACCAGGGTCCTCGGGACTTCCAGACACTGGAAAACTGGATGC
TGCAGACACTGAACGAGGAGCCAGTGACACCAGAGCCGGAAGTGGAACCGCCCAGTGCCCCCGAGCTCAAGCAAG
GGCTGTATGAGCTCTCAGCAAGCAACTTTGAGCTGCACGTTGCACAAGGCGACCACTTTATCAAGTTCTTCGCTC
CGTGGTGTGGTCACTGCAAAGCCCTGGCTCCAACCTGGGAGCAGCTGGCTCTGGGCCTTGAACATTCCGAAACTG
TCAAGATTGGCAAGGTTGATTGTACACAGCACTATGAACTCTGCTCCGGAAACCAGGTTCGTGGCTATCCCACTC
TTCTCTGGTTCCGAGATGGGAAAAAGGTGGATCAGTACAAGGGAAAGCGGGATTTGGAGTCACTGAGGGAGTACG
TGGAGTCGCAGCTGCAGCGCACAGAGACTGGAGCGACGGAGACCGTCACGCCCTCAGAGGCCCCGGTGCTGGCAG
CTGAGCCCGAGGCTGACAAGGGCACTGTGTTGGCACTCACTGAAAATAACTTCGATGACACCATTGCAGAAGGAA
TAACCTTCATCAAGTTTTATGCTCCATGGTGTGGTCATTGTAAGACTCTGGCTCCTACTTGGGAGGAACTCTCTA
AAAAGGAATTCCCTGGTCTGGCGGGGGTCAAGATCGCCGAAGTAGACTGCACTGCTGAACGGAATATCTGCAGCA
AGTATTCGGTACGAGGCTACCCCACGTTATTGCTTTTCCGAGGAGGGAAGAAAGTCAGTGAGCACAGTGGAGGCA
GAGACCTTGACTCGTTACACCGCTTTGTCCTGAGCCAAGCGAAAGACGAAC**TTTAG**GAACACAGTTGGAGGTCAC
CTCTCCTGCCCAGCTCCCGCACCCTGCGTTTAGGAGTTCAGTCCCACAGAGGCCACTGGGTTCCCAGTGGTGGCT
GTTCAGAAAGCAGAACATACTAAGCGTGAGGTATCTTCTTTGTGTGTGTGTTTTCCAAGCCAACACACTCTACAG
ATTCTTTATTAAGTTAAGTTTCTCTAAGTAAATGTGTAACTCATGGTCACTGTGTAAACATTTTCAGTGGCGATA
TATCCCCTTTGACCTTCTCTTGATGAAATTTACATGGTTTCCTTTGAGACTAAAATAGCGTTGAGGGAAATGAAA
TTGCTGGACTATTTGTGGCTCCTGAGTTGAGTGATTTTGGTGAAAGAAAGCACATCCAAAGCATAGTTTACCTGC
CCACGAGTTCTGGAAAGGTGGCCTTGTGGCAGTATTGACGTTCCTCTGATCTTAAGGTCACAGTTGACTCAATAC
TGTGTTGGTCCGTAGCATGGAGCAGATTGAAATGCAAAAACCCACACCTCTGGAAGATACCTTCACGGCCGCTGC
TGGAGCTTCTGTTGCTGTGAATACTTCTCTCAGTGTGAGAGGTTAGCCGTGATGAAAGCAGCGTTACTTCTGACC
GTGCCTGAGTAAGAGAATGCTGATGCCATAACTTTATGTGTCGATACTTGTCAAATCAGTTACTGTTCAGGGGAT
CCTTCTGTTTCTCACGGGGTGAAACATGTCTTTAGTTCCTCATGTTAACACGAAGCCAGAGCCCACATGAACTGT
TGGATGTCTTCCTTAGAAAGGGTAGGCATGGAAAATTCCACGAGGCTCATTCTCAGTATCTCATTAACTCATTGA
AAGATTCCAGTTGTATTTGTCACCTGGGGTGACAAGACCAGACAGGCTTTCCCAGGCCTGGGTATCCAGGGAGGC
TCTGCAGCCCTGCTGAAGGGCCCTAACTAGAGTTCTAGAGTTTCTGATTCTGTTTCTCAGTAGTCCTTTTAGAGG
CTTGCTATACTTGGTCTGCTTCAAGGAGGTCGACCTTCTAATGTATGAAGAATGGGATGCATTTGATCTCAAGAC
CAAAGACAGATGTCAGTGGGCTGCTCTGGCCCTGGTGTGCACGGCTGTGGCAGCTGTTGATGCCAGTGTCCTCTA
ACTCATGCTGTCCTTGTGATTAAACACCTCTATCTCCCTTGGGAATAAGCACATACAGGCTTAAGCTCTAAGATA
GATAGGTGTTTGTCCTTTTACCATCGAGCTACTTCCCATAATAACCACTTTGCATCCAACACTCTTCACCCACCT
CCCATACGCAAGGGGATGTGGATACTTGGCCCAAAGTAACTGGTGGTAGGAATCTTAGAAACAAGACCACTTATA
CTGTCTGTCTGAGGCAGAAGATAACAGCAGCATCTCGACCAGCCTCTGCCTTAAAGGAAATCTTTATTAATCACG
TATGGTTCACAGATAATTCTTTTTTTAAAAAAACCCAACCTCCTAGAGAAGCACAACTGTCAAGAGTCTTGTACA
CACAACTTCAGCTTTGCATCACGAGTCTTGTATTCCAAGAAAATCAAAGTGGTACAATTTGTTTGTTTACACTAT
GATACTTTCTAAATAAACTCTTTTTTTTTAA

FIGURE 35

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776

><subunit 1 of 1, 432 aa, 1 stop

><MW: 47629, pI: 5.90, NX(S/T): 0

MPARPGRLLPLLARPAALTALLLLLLGHGGGGRWGARAQEAAAAAADGPPAADGEDGQDPHS

KHLYTADMFTHGIQSAAHFVMFFAPWCGHCQRLQPTWNDLGDKYNSMEDAKVYVAKVDCTAH

SDVCSAQGVRGYPTLKLFKPGQEAVKYQGPRDFQTLENWMLQTLNEEPVTPEPEVEPPSAPE

LKQGLYELSASNFELHVAQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCT

QHYELCSGNQVRGYPTLLWFRDGKKVDQYKGKRDLESLREYVESQLQRTETGATETVTPSEA

PVLAAEPEADKGTVLALTENNFDDTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA

GVKIAEVDCTAERNICSKYSVRGYPTLLLFRGGKKVSEHSGGRDLDSLHRFVLSQAKDEL

Signal sequence:

amino acids 1-32

FIGURE 36

CTTTTCTGAGGAACCACAGCA**ATG**AATGGCTTTGCATCCTTGCTTCGAAGAAACCAATTTAT

CCTCCTGGTACTATTTCTTTTGCAAATTCAGAGTCTGGGTCTGGATATTGATAGCCGTCCTA

CCGCTGAAGTCTGTGCCACACACACAATTTCACCAGGACCCAAAGGAGATGATGGTGAAAAA

GGAGATCCAGGAGAAGAGGGAAAGCATGGCAAAGTGGGACGCATGGGGCCGAAAGGAATTAA

AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCCATTGGGAAGA

AGGGTGACAAAGGGGAAAAAGGTTTGCTTGGAATACCTGGAGAAAAAGGCAAAGCAGGTACT

GTCTGTGATTGTGGAAGATACCGGAAATTTGTTGGACAACTGGATATTAGTATTGCTCGGCT

CAAGACATCTATGAAGTTTGTCAAGAATGTGATAGCAGGGATTAGGGAAACTGAAGAGAAAT

TCTACTACATCGTGCAGGAAGAGAAGAACTACAGGGAATCCCTAACCCACTGCAGGATTCGG

GGTGGAATGCTAGCCATGCCCAAGGATGAAGCTGCCAACACACTCATCGCTGACTATGTTGC

CAAGAGTGGCTTCTTTCGGGTGTTCATTGGCGTGAATGACCTTGAAAGGGAGGGACAGTACA

TGTCCACAGACAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGGAACCCAGCGAC

CCCTATGGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTG

CCATCTTACCATGTACTTTGTCTGTGAGTTCATCAAGAAGAAAAAG**TAA**CTTCCCTCATCCT

ACGTATTTGCTATTTTCCTGTGACCGTCATTACAGTTATTGTTATCCATCCTTTTTTTCCTG

ATTGTACTACATTTGATCTGAGTCAACATAGCTAGAAAATGCTAAACTGAGGTATGGAGCCT

CCATCATCAAAAAAAAAAAAAAAAA

FIGURE 37

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980

><subunit 1 of 1, 277 aa, 1 stop

><MW: 30645, pI: 7.47, NX(S/T): 2

MNGFASLLRRNQFILLVLFLLQIQSLGLDIDSRPTAEVCATHTISPGPKGDDGEKGDPGEEG

KHGKVGRMGPKGIKGELGDMGDQGNIGKTGPIGKKGDKGEKGLLGIPGEKGKAGTVCDCGRY

RKFVGQLDISIARLKTSMKFVKNVIAGIRETEEKFYYIVQEEKNYRESLTHCRIRGGMLAMP

KDEAANTLIADYVAKSGFFRVFIGVNDLEREGQYMSTDNTPLQNYSNWNEGEPSDPYGHEDC

VEMLSSGRWNDTECHLTMYFVCEFIKKKK

Signal peptide:

amino acids 1-25

FIGURE 38

GGTTCTATCGATTCGAATTCGGCCACACTGGCCGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCTGCTCTCCGCCCGTGTGGAGTGGTGGGGGCCTGGGTGGGA**ATG**GGCGTGT
GCCAGCGCACGCGCGCTCCCTGGAAGGAGAAGTCTCAGCTAGAACGAGCGGCCCTAGGTTTT
CGGAAGGGAGGATCAGGGATGTTTGCGAGCGGCTGGAACCAGACGGTGCCGATAGAGGAAGC
GGGCTCCATGGCTGCCCTCCTGCTGCTGCCCCTGCTGCTGTTGCTACCGCTGCTGCTGCTGA
AGCTACACCTCTGGCCGCAGTTGCGCTGGCTTCCGGCGGACTTGGCCTTTGCGGTGCGAGCT
CTGTGCTGCAAAAGGGCTCTTCGAGCTCGCGCCCTGGCCGCGGCTGCCGCCGACCCGGAAGG
TCCCGAGGGGGGCTGCAGCCTGGCCTGGCGCCTCGCGGAACTGGCCCAGCAGCGCGCCGCGC
ACACCTTTCTCATTCACGGCTCGCGGCGCTTTAGCTACTCAGAGGCGGAGCGCGAGAGTAAC
AGGGCTGCACGCGCCTTCCTACGTGCGCTAGGCTGGGACTGGGGACCCGACGGCGGCGACAG
CGGCGAGGGGAGCGCTGGAGAAGGCGAGCGGGCAGCGCCGGGAGCCGGAGATGCAGCGGCCG
GAAGCGGCGCGGAGTTTGCCGGAGGGGACGGTGCCGCCAGAGGTGGAGGAGCCGCCGCCCCT
CTGTCACCTGGAGCAACTGTGGCGCTGCTCCTCCCCGCTGGCCCAGAGTTTCTGTGGCTCTG
GTTCGGGCTGGCCAAGGCCGGCCTGCGCACTGCCTTTGTGCCCACCGCCCTGCGCCGGGGCC
CCCTGCTGCACTGCCTCCGCAGCTGCGGCGCGCGCGCTGGTGCTGGCGCCAGAGTTTCTG
GAGTCCCTGGAGCCGGACCTGCCCGCCCTGAGAGCCATGGGGCTCCACCTGTGGGCTGCAGG
CCCAGGAACCCACCCTGCTGGAATTAGCGATTTGCTGGCTGAAGTGTCCGCTGAAGTGGATG
GGCCAGTGCCAGGATACCTCTCTTCCCCCCAGAGCATAACAGACACGTGCCTGTACATCTTC
ACCTCTGGCACCACGGGCCTCCCCAAGGCTGCTCGGATCAGTCATCTGAAGATCCTGCAATG
CCAGGGCTTCTATCAGCTGTGTGGTGTCCACCAGGAAGATGTGATCTACCTCGCCCTCCCAC
TCTACCACATGTCCGGTTCCCTGCTGGGCATCGTGGGCTGCATGGGCATTGGGGCCACAGTG
GTGCTGAAATCCAAGTTCTCGGCTGGTCAGTTCTGGGAAGATTGCCAGCAGCACAGGGTGAC
GGTGTTCCAGTACATTGGGGAGCTGTGCCGATACCTTGTCAACCAGCCCCCGAGCAAGGCAG
AACGTGGCCATAAGGTCCGGCTGGCAGTGGGCAGCGGGCTGCGCCCAGATACCTGGGAGCGT
TTTGTGCGGCGCTTCGGGCCCCTGCAGGTGCTGGAGACATATGGACTGACAGAGGGCAACGT
GGCCACCATCAACTACACAGGACAGCGGGGCGCTGTGGGGCGTGCTTCCTGGCTTTACAAGC
ATATCTTCCCCTTCTCCTTGATTCGCTATGATGTCACCACAGGAGAGCCAATTCGGGACCCC
CAGGGGCACTGTATGGCCACATCTCCAGGTGAGCCAGGGCTGCTGGTGGCCCCGGTAAGCCA
GCAGTCCCCATTCCTGGGCTATGCTGGCGGGCCAGAGCTGGCCCAGGGGAAGTTGCTAAAGG
ATGTCTTCCGGCCTGGGGATGTTTTCTTCAACACTGGGGACCTGCTGGTCTGCGATGACCAA
GGTTTTCTCCGCTTCCATGATCGTACTGGAGACACCTTCAGGTGGAAGGGGGAGAATGTGGC
CACAACCGAGGTGGCAGAGGTCTTCGAGGCCCTAGATTTTCTTCAGGAGGTGAACGTCTATG
GAGTCACTGTGCCAGGGCATGAAGGCAGGGCTGGAATGGCAGCCCTAGTTCTGCGTCCCCCC
CACGCTTTGGACCTTATGCAGCTCTACACCCACGTGTCTGAGAACTTGCCACCTTATGCCCG
GCCCCGATTCCTCAGGCTCCAGGAGTCTTTGGCCACCACAGAGACCTTCAAACAGCAGAAAG
TTCGGATGGCAAATGAGGGCTTCGACCCCAGCACCCTGTCTGACCCACTGTACGTTCTGGAC
CAGGCTGTAGGTGCCTACCTGCCCCTCACAACTGCCCGGTACAGCGCCCTCCTGGCAGGAAA
CCTTCGAATC**TGA**GAACTTCCACACCTGAGGCACCTGAGAGAGGAACTCTGTGGGGTGGGGG
CCGTTGCAGGTGTACTGGGCTGTCAGGGATCTTTTCTATACCAGAACTGCGGTCACTATTTT
GTAATAAATGTGGCTGGAGCTGATCCAGCTGTCTCTGACCTAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGTAGGGATAACAGGGTAATAAGC
TTGGCCGCCATGGCCCAACTTGTTTATTGCAG

FIGURE 39

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913

><subunit 1 of 1, 730 aa, 1 stop

><MW: 78644, pI: 7.65, NX(S/T): 2

MGVCQRTRAPWKEKSQLERAALGFRKGGSGMFASGWNQTVPIEEAGSMAALLLLPLLLLLLPL

LLLKLHLWPQLRWLPADLAFAVRALCCKRALRARALAAAAADPEGPEGGCSLAWRLAELAQQ

RAAHTFLIHGSRRFSYSEAERESNRAARAFLRALGWDWGPDGGDSGEGSAGEGERAAPGAGD

AAAGSGAEFAGGDGAARGGGAAAPLSPGATVALLLPAGPEFLWLWFGLAKAGLRTAFVPTAL

RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLLAEVSA

EVDGPVPGYLSSPQSITDTCLYIFTSGTTGLPKAARISHLKILQCQGFYQLCGVHQEDVIYL

ALPLYHMSGSLLGIVGCMGIGATVVLKSKFSAGQFWEDCQQHRVTVFQYIGELCRYLVNQPP

SKAERGHKVRLAVGSGLRPDTWERFVRRFGPLQVLETYGLTEGNVATINYTGQRGAVGRASW

LYKHIFPFSLIRYDVTTGEPIRDPQGHCMATSPGEPGLLVAPVSQQSPFLGYAGGPELAQGK

LLKDVFRPGDVFFNTGDLLVCDDQGFLRFHDRTGDTFRWKGENVATTEVAEVFEALDFLQEV

NVYGVTVPGHEGRAGMAALVLRPPHALDLMQLYTHVSENLPPYARPRFLRLQESLATTETFK

QQKVRMANEGFDPSTLSDPLYVLDQAVGAYLPLTTARYSALLAGNLRI

Type II transmembrane domain:

amino acids 45-65

Other transmembrane domain:

amino acids 379-398

cAMP- and cGMP-dependent protein kinase phosphorylation site

starting at amino acid 136

CUB domain protein motif

amino acids 254-261

putative AMP-binding domain siganture

amino acids 332-343

N-glycosylation sites

amino acids 37-40 and 483-486

FIGURE 40

CCTGTGTTAAGCTGAGGTTTCCCCTAGATCTCGTATATCCCCAACACATACCTCCACGCACA
CACATCCCCAAGAACCTCGAGCTCACACCAACAGACACACGCGCGCATACACACTCGCTCTC
GCTTGTCCATCTCCCTCCCGGGGGAGCCGGCGCGCGCTCCCACCTTTGCCGCACACTCCGGC
GAGCCGAGCCCGCAGCGCTCCAGGATTCTGCGGCTCGGAACTCGGATTGCAGCTCTGAACCC
CCATGGTGGTTTTTTAAACACTTCTTTTCCTTCTCTTCCTCGTTTTGATTGCACCGTTTCCA
TCTGGGGGCTAGAGGAGCAAGGCAGCAGCCTTCCCAGCCAGCCCTTGTTGGCTTGCCATCGT
CCATCTGGCTTATAAAAGTTTGCTGAGCGCAGTCCAGAGGGCTGCGCTGCTCGTCCCCTCGG
CTGGCAGAAGGGGGTGACGCTGGGCAGCGGCGAGGAGCGCGCCGCTGCCTCTGGCGGGCTTT
CGGCTTGAGGGGCAAGGTGAAGAGCGCACCGGCCGTGGGGTTTACCGAGCTGGATTTGTATG
TTGCACC**ATG**CCTTCTTGGATCGGGGCTGTGATTCTTCCCCTCTTGGGGCTGCTGCTCTCCC
TCCCCGCCGGGGCGGATGTGAAGGCTCGGAGCTGCGGAGAGGTCCGCCAGGCGTACGGTGCC
AAGGGATTCAGCCTGGCGGACATCCCCTACCAGGAGATCGCAGGGGAACACTTAAGAATCTG
TCCTCAGGAATATACATGCTGCACCACAGAAATGGAAGACAAGTTAAGCCAACAAAGCAAAC
TCGAATTTGAAAACCTTGTGGAAGAGACAAGCCATTTTGTGCGCACCACTTTTGTGTCCAGG
CATAAGAAATTTGACGAATTTTTCCGAGAGCTCCTGGAGAATGCAGAAAAGTCACTAAATGA
TATGTTTGTACGGACCTATGGCATGCTGTACATGCAGAATTCAGAAGTCTTCCAGGACCTCT
TCACAGAGCTGAAAAGGTACTACACTGGGGGTAATGTGAATCTGGAGGAAATGCTCAATGAC
TTTTGGGCTCGGCTCCTGGAACGGATGTTTCAGCTGATAAACCCTCAGTATCACTTCAGTGA
AGACTACCTGGAATGTGTGAGCAAATACACTGACCAGCTCAAGCCATTTGGAGACGTGCCCC
GGAAACTGAAGATTCAGGTTACCCGCGCCTTCATTGCTGCCAGGACCTTTGTCCAGGGGCTG
ACTGTGGGCAGAGAAGTTGCAAACCGAGTTTCCAAGGTCAGCCCAACCCCAGGGTGTATCCG
TGCCCTCATGAAGATGCTGTACTGCCCATACTGTCGGGGGCTTCCCACTGTGAGGCCCTGCA
ACAACTACTGTCTCAACGTCATGAAGGGCTGCTTGGCAAATCAGGCTGACCTCGACACAGAG
TGGAATCTGTTTATAGATGCAATGCTCTTGGTGGCAGAGCGACTGGAGGGGCCATTCAACAT
TGAGTCGGTCATGGACCCGATAGATGTCAAGATTTCTGAAGCCATTATGAACATGCAAGAAA
ACAGCATGCAGGTGTCTGCAAAGGTCTTTCAGGGATGTGGTCAGCCCAAACCTGCTCCAGCC
CTCAGATCTGCCCGCTCAGCTCCTGAAAATTTTAATACACGTTTCAGGCCCTACAATCCTGA
GGAAAGACCAACAACTGCTGCAGGCACAAGCTTGGACCGGCTGGTCACAGACATAAAAGAGA
AATTGAAGCTCTCTAAAAAGGTCTGGTCAGCATTACCCTACACTATCTGCAAGGACGAGAGC
GTGACAGCGGGCACGTCCAACGAGGAGGAATGCTGGAACGGGCACAGCAAAGCCAGATACTT
GCCTGAGATCATGAATGATGGGCTCACCAACCAGATCAACAATCCCGAGGTGGATGTGGACA
TCACTCGGCCTGACACTTTCATCAGACAGCAGATTATGGCTCTCCGTGTGATGACCAACAAA
CTAAAAAACGCCTACAATGGCAATGATGTCAATTTCCAGGACACAAGTGATGAATCCAGTGG
CTCAGGGAGTGGCAGTGGGTGCATGGATGACGTGTGTCCCACGGAGTTTGAGTTTGTCACCA
CAGAGGCCCCCGCAGTGGATCCCGACCGGAGAGAGGTGGACTCTTCTGCAGCCCAGCGTGGC
CACTCCCTGCTCTCCTGGTCTCTCACCTGCATTGTCCTGGCACTGCAGAGACTGTGCAGA**TA**
ATCTTGGGTTTTTGGTCAGATGAAACTGCATTTTAGCTATCTGAATGGCCAACTCACTTCTT
TTCTTACACTCTTGGACAATGGACCATGCCACAAAAACTTACCGTTTTCTATGAGAAGAGAG
CAGTAATGCAATCTGCCTCCCTTTTTGTTTTCCCAAAGAGTACCGGGTGCCAGACTGAACTG
CTTCCTCTTTCCTTCAGCTATCTGTGGGGACCTTGTTTATTCTAGAGAGAATTCTTACTCAA
ATTTTTCGTACCAGGAGATTTTCTTACCTTCATTTGCTTTTATGCTGCAGAAGTAAAGGAAT
CTCACGTTGTGAGGGTTTTTTTTTTCTCATTTAAAAT

FIGURE 41

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914

><subunit 1 of 1, 555 aa, 1 stop

><MW: 62736, pI: 5.36, NX(S/T): 0

MPSWIGAVILPLLGLLLSLPAGADVKARSCGEVRQAYGAKGFSLADIPYQEIAGEHLRICPQ

EYTCCTTEMEDKLSQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELLENAEKSLNDMF

VRTYGMLYMQNSEVFQDLFTELKRYYTGGNVNLEEMLNDFWARLLERMFQLINPQYHFSEDY

LECVSKYTDQLKPFGDVPRKLKIQVTRAFIAARTFVQGLTVGREVANRVSKVSPTPGCIRAL

MKMLYCPYCRGLPTVRPCNNYCLNVMKGCLANQADLDTEWNLFIDAMLLVAERLEGPFNIES

VMDPIDVKISEAIMNMQENSMQVSAKVFQGCGQPKPAPALRSARSAPENFNTRFRPYNPEER

PTTAAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDESVTAGTSNEEECWNGHSKARYLPE

IMNDGLTNQINNPEVDVDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQDTSDESSGSG

SGSGCMDDVCPTEFEFVTTEAPAVDPDRREVDSSAAQRGHSLLSWSLTCIVLALQRLCR

Signal peptide:

amino acids 1-23

FIGURE 42A

CGGACGCGTGGGCGGACGCGTGGGCAAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCTGAGAAAACGCACG
CAGTTTGCAGCGCCTGCGCCGGGTGCGCCAACTACGCAAAGACCAAGCGGGCTCCGCGCGGACCGGCCGCGGGGC
TAGGGACCCGGCTTTGGCCTTCAGGCTCCCTAGCAGCGGGGAAAAGGAATTGCTGCCCGGAGTTTCTGCGGAGGT
GGAGGGAGATCAGGAAACGGCTTCTTCCTCACTTCGCCGCCTGGTGAGTGTCGGGGAGATTGGCAAACGCCTAGG
AAAGGACTGGGGAAAATAGCCCTGGGAAAGTGGAGAAGGTGATCAGGAGGCCGGTCCACTACGGCAGTTTATCTG
TCTGATCAGAGCCAGACGCGACGCGTCCACTTCGCAGTTCTTTCCAGGTGTGGGGACCGCAGGACAGACGGCCGA
TCCCGCCGCCCTCCGTACCAGCACTCCCAGGAGAGTCAGCCTCGCTCCCCAACGTCGAGGGCGCTCTGGCCACGA
AAAGTTCCTGTCCACTGTGATTCTCAATTCCTTGCTTGGTTTTTTTCTCCAGAGAACTTTTGGGTGGAGATATTA
ACTTTTTTCTTTTTTTTTTTCCTTGGTGGAAGCTGCTCTAGGGAGGGGGGAGGAGGAGGAGAAAGTGAAATGTGC
TGGAGAAGAGCGAGCCCTCCTTGTTCTTCCGGAGTCCCATCCATTAAGCCATCACTTCTGGAAGATTAAAGTTGT
CGGACATGGTGACAGCTGAGAGGAGAGGAGGATTTCTTGCCAGGTGGAGAGTCTTCACCGTCTGTTGGGTGCATG
TGTGCGCCCGCAGCGGCGCGGGGCGCGTGGTTCTCCGCGTGGAGTCTCACCTGGGACCTGAGTGA**ATG**GCTCCCA
GGGGCTGTGCGGGGCATCCGCCTCCGCCTTCTCCACAGGCCTGTGTCTGTCCTGGAAAGATGCTAGCAATGGGGG
CGCTGGCAGGATTCTGGATCCTCTGCCTCCTCACTTATGGTTACCTGTCCTGGGGCCAGGCCTTAGAAGAGGAGG
AAGAAGGGGCCTTACTAGCTCAAGCTGGAGAGAAACTAGAGCCCAGCACAACTTCCACCTCCCAGCCCCATCTCA
TTTTCATCCTAGCGGATGATCAGGGATTTAGAGATGTGGGTTACCACGGATCTGAGATTAAAACACCTACTCTTG
ACAAGCTCGCTGCCGAAGGAGTTAAACTGGAGAACTACTATGTCCAGCCTATTTGCACACCATCCAGGAGTCAGT
TTATTACTGGAAAGTATCAGATACACACCGGACTTCAACATTCTATCATAAGACCTACCCAACCCAACTGTTTAC
CTCTGGACAATGCCACCCTACCTCAGAAACTGAAGGAGGTTGGATATTCAACGCATATGGTCGGAAAATGGCACT
TGGGTTTTAACAGAAAAGAATGCATGCCCACCAGAAGAGGATTTGATACCTTTTTTGGTTCCCTTTTGGGAAGTG
GGGATTACTATACACTACAAATGTGACAGTCCTGGGATGTGTGGCTATGACTTGTATGAAAACGACAATGCTG
CCTGGGACTATGACAATGGCATATACTCCACACAGATGTACACTCAGAGAGTACAGCAAATCTTAGCTTCCCATA
ACCCCACAAAGCCTATATTTTTATATACTGCCTATCAAGCTGTTCATTCACCACTGCAAGCTCCTGGCAGGTATT
TCGAACACTACCGATCCATTATCAACATAAACAGGAGAAGATATGCTGCCATGCTTTCCTGCTTAGATGAAGCAA
TCAACAACGTGACATTGGCTCTAAAGACTTATGGTTTCTATAACAACAGCATTATCATTTACTCTTCAGATAATG
GTGGCCAGCCTACGGCAGGAGGGAGTAACTGGCCTCTCAGAGGTAGCAAAGGAACATATTGGGAAGGAGGGATCC
GGGCTGTAGGCTTTGTGCATAGCCCACTTCTGAAAAACAAGGGAACAGTGTGTAAGGAACTTGTGCACATCACTG
ACTGGTACCCCACTCTCATTTCACTGGCTGAAGGACAGATTGATGAGGACATTCAACTAGATGGCTATGATATCT
GGGAGACCATAAGTGAGGGTCTTCGCTCACCCCGAGTAGATATTTTGCATAACATTGACCCCTATACACCAAGGC
AAAAAATGGCTCCTGGGCAGCAGGCTATGGGATCTGGAACACTGCAATCCAGTCAGCCATCAGAGTGCAGCACTG
GAAATTGCTTACAGGAAATCCTGGCTACAGCGACTGGGTCCCCCCTCAGTCTTTCAGCAACCTGGGACCGAACCG
GTGGCACAATGAACGGATCACCTTGTCAACTGGCAAAAGTGTATGGCTTTTCAACATCACAGCCGACCCATATGA
GAGGGTGGACCTATCTAACAGGTATCCAGGAATCG**TGA**AGAAGCTCCTACGGAGGCTCTCACAGTTCAACAAAAC
TGCAGTGCCGGTCAGGTATCCCCCCAAAGACCCCAGAAGTAACCCTAGGCTCAATGGAGGGGTCTGGGGACCATG
GTATAAAGAGGAAACCAAGAAAAAGAAGCCAAGCAAAAATCAGGCTGAGAAAAAGCAAAGAAAAGCAAAAAAAA
GAAGAAGAAACAGCAGAAAGCAGTCTCAGGTAAACCAGCAAATTTGGCTCGATAATATCGCTGGCCTAAGCGTCA
·GGCTTGTTTTCATGCTGTGCCACTCCAGAGACTTCTGCCACCTGGCCGCCACACTGAAAACTGTCCTGCTCAGTG
CCAAGGTGCTACTCTTGCAAGCCACACTTAGAGAGAGTGGAGATGTTTATTTCTCTCGCTCCTTTAGAAAACGTG
GTGAGTCCTGAGTTCCACTGCTGTGCTTCAGTCAACTGACCAAACACTGCTTTGAATTATAGGAGGAGAACAATA
ACCTACCATCCGCAAGCATGCTAATTTGATGGAAGTTACAGGGTAGCATGATTAAAACTACCTTTGATAAATTAC

FIGURE 42B

AGTCAAAGATTGTGTCACCTCAAAGGCCTTGAAGAATATATTTTCTTGGTGAATTTTTGTATGTCTGTCATATGA

CACTTGGGTTTTTTAATTAATTCTATTTTATATATATAAATATATGTTTCTTTTCCTGTGAAAAGCTGTTTTTCT

CACATGTGAACAGCTTGCACCTCATTTTACCATGCGTGAGGGAATGGCAAATAAGAATGTTTGAGCACACTGCCC

ACAATGAATGTAACTATTTTCTAAACACTTTACTAGAAGAACATTTCAGTATAAAAAACCTAATTTATTTTTACA

GAAAATATTTTGTTGTTTTTATAAAAAGTTATGCAAATGACTTTTATTTTTATTTCCTGCATACCATTAGAAGA

ATTTTATTTCATTTCTTCAAATTATCAAGCACTGTAATACTATAAATTAATGTAATACTGTGTGAATTCAGACTA

TAAAAAACATCATTCAGAAAACTTTATAATCGTCATTGTTCAATCAAGATTTTGAATGTAATAAGATGAATATAT

ATTACTTGGAAATTCAATGTTTGTGCAGAGTTGAGACAACTTTATTGTTTCTATCATAAACTATTTATGTATCTT

AATTATTAAAATGATTTACTTTATGGCACTAGAAAATTTACTGTGGCTTTTCTGATCTAACTTCTAGCTAAAATT

GTATCATTGGTCCTAAAAAATAAAAATCTTTACTAATAGGCAATTGAAGGAATGGTTTGCTAACAACCACAGTAA

TATAATATGATTTTACAGATAGATGCTTCCCCTTGGCTATGACATGGAGAAAGATTTTCCCATAATAATAACTAA

TATTTATATTAGGTTGGTGCAAAACTAGTTGCGGTTTTTCCCATTAAAAGTAATAACCTTACTCTTATACAAAGT

GGACACTGTGGGGAGATACAGAGAAATGGAAGATACGGATCCTGCCTGGAGTAGGTAACCTTGCTTGGAAACCCC

ACATGCAAACGTCATGAGGAGAATTAAAGGAGTATTATCAGTAATGAAGTTTATCATGGGTCATCAATGAGCATA

GATTGGTGTGGATCCTGTAGACCCTGGTGTTTTCTTTGAAGTGCCCTCTCCTAATGCAGAGGCCTTGAAGCTTAC

AGTATACACTTGAAAAGTCACAGATAGCTAGAATTATGATCTTTGAAGTTATAACTGTGATCTGAAAATGTGTGT

GGTGGTATGACAGCATACCATTAAATACATTTACATCACAGCTCAAAGGACTGTGATATAATCCATTTATATCAC

AACTCAAAGGACTGTGATATAATCCATTTATATCACAGCTCACAGTTTCTGAAAATGTATAAAAGAATCTATAAT

CTAGTACTGAAATTACTAAATTGGGTAAGATGATTTAAATGATTTTAATTTTAACATTTTATTTCTAGAATATAT

GGCTCCATTTTATTTTATAGTGTAAAGTTGTATTTCCTAAAGTTTGTGTTTTGTCGACAGTATCTTTTAAATGAG

TCTTAAAAATAAAGGCATATTGTTCATGTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 43

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296

><subunit 1 of 1, 515 aa, 1 stop

><MW: 56885, pI: 6.49, NX(S/T): 5

MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLAQA
GEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTP
SRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFNRKEC
MPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQ
ILASHNPTKPIFLYTAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLA
LKTYGFYNNSIIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCK
ELVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPYTPRQKMAPG
QQAMGSGTLQSSQPSECSTGNCLQEILATATGSPLSLSATWDRTGGTMNGSPCQLAKVYGFS
TSQPTHMRGWTYLTGIQES

Important Features:

Signal Peptide:
amino acids 1-37

Sulfatases signature 1.
amino acids 120-132

Sulfatases signature 2.
amino acids 168-177

Tyrosine kinase phosphorylation site.
amino acids 163-169

N-glycosylation sites.
amino acids 157-160, 306-309 and 318-321

FIGURE 44

```
CGGACGCGTGGGTGCGAGTGGAGCGGAGGACCCGAGCGGCTGAGGAGAGAGGAGGCGGCGGC
TTAGCTGCTACGGGGTCCGGCCGGCGCCCTCCCGAGGGGGGCTCAGGAGGAGGAAGGAGGAC
CCGTGCGAGAATGCCTCTGCCCTGGAGCCTTGCGCTCCCGCTGCTGCTCTCCTGGGTGGCAG
GTGGTTTCGGGAACGCGGCCAGTGCAAGGCATCACGGGTTGTTAGCATCGGCACGTCAGCCT
GGGGTCTGTCACTATGGAACTAAACTGGCCTGCTGCTACGGCTGGAGAAGAAACAGCAAGGG
AGTCTGTGAAGCTACATGCGAACCTGGATGTAAGTTTGGTGAGTGCGTGGGACCAAACAAAT
GCAGATGCTTTCCAGGATACACCGGGAAAACCTGCAGTCAAGATGTGAATGAGTGTGGAATG
AAACCCCGGCCATGCCAACACAGATGTGTGAATACACACGGAAGCTACAAGTGCTTTTGCCT
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGTGAACTCTAGGACATGTGCCATGATAA
ACTGTCAGTACAGCTGTGAAGACACAGAAGAAGGGCCACAGTGCCTGTGTCCATCCTCAGGA
CTCCGCCTGGCCCCAAATGGAAGAGACTGTCTAGATATTGATGAATGTGCCTCTGGTAAAGT
CATCTGTCCCTACAATCGAAGATGTGTGAACACATTTGGAAGCTACTACTGCAAATGTCACA
TTGGTTTCGAACTGCAATATATCAGTGGACGATATGACTGTATAGATATAAATGAATGTACT
ATGGATAGCCATACGTGCAGCCACCATGCCAATTGCTTCAATACCCAAGGGTCCTTCAAGTG
TAAATGCAAGCAGGGATATAAAGGCAATGGACTTCGGTGTTCTGCTATCCCTGAAAATTCTG
TGAAGGAAGTCCTCAGAGCACCTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGCTCAC
AAAAACAGCATGAAAAAGAAGGCAAAAATTAAAAATGTTACCCCAGAACCCACCAGGACTCC
TACCCCTAAGGTGAACTTGCAGCCCTTCAACTATGAAGAGATAGTTTCCAGAGGCGGGAACT
CTCATGGAGGTAAAAAAGGGAATGAAGAGAAATGAAAGAGGGGCTTGAGGATGAGAAAAGAG
AAGAGAAAGCCCTGAAGAATGACATAGAGGAGCGAAGCCTGCGAGGAGATGTGTTTTTCCCT
AAGGTGAATGAAGCAGGTGAATTCGGCCTGATTCTGGTCCAAAGGAAAGCGCTAACTTCCAA
ACTGGAACATAAAGATTTAAATATCTCGGTTGACTGCAGCTTCAATCATGGGATCTGTGACT
GGAAACAGGATAGAGAAGATGATTTTGACTGGAATCCTGCTGATCGAGATAATGCTATTGGC
TTCTATATGGCAGTTCCGGCCTTGGCAGGTCACAAGAAAGACATTGGCCGATTGAAACTTCT
CCTACCTGACCTGCAACCCCAAAGCAACTTCTGTTTGCTCTTTGATTACCGGCTGGCCGGAG
ACAAAGTCGGGAAACTTCGAGTGTTTGTGAAAAACAGTAACAATGCCCTGGCATGGGAGAAG
ACCACGAGTGAGGATGAAAAGTGGAAGACAGGGAAAATTCAGTTGTATCAAGGAACTGATGC
TACCAAAAGCATCATTTTTGAAGCAGAACGTGGCAAGGGCAAAACCGGCGAAATCGCAGTGG
ATGGCGTCTTGCTTGTTTCAGGCTTATGTCCAGATAGCCTTTTATCTGTGGATGACTGAATG
TTACTATCTTTATATTTGACTTTGTATGTCAGTTCCCTGGTTTTTTTGATATTGCATCATAG
GACCTCTGGCATTTTAGAATTACTAGCTGAAAAATTGTAATGTACCAACAGAAATATTATTG
TAAGATGCCTTTCTTGTATAAGATATGCCAATATTTGCTTTAAATATCATATCACTGTATCT
TCTCAGTCATTTCTGAATCTTTCCNCATTATATTATAAAATNTGGAAANGTCAGTTTATCTC
CCCTCCTCNGTATATCTGATTTGTATANGTANGTTGATGNGCTTCTCTCTACAACATTTCTA
GAAAATAGAAAAAAAAGCACAGAGAAATGTTTAACTGTTTGACTCTTATGATACTTCTTGGA
AACTATGACATCAAAGATAGACTTTTGCCTAAGTGGCTTAGCTGGGTCTTTCATAGCCAAAC
TTGTATATTTAATTCTTTGTAATAATAA
```

FIGURE 45

MPLPWSLALPLLLSWVAGGFGNAASARHHGLLASARQPGVCHYGTKLACCYGWRRNSKGVCE
ATCEPGCKFGECVGPNKCRCFPGYTGKTCSQDVNECGMKPRPCQHRCVNTHGSYKCFCLSGH
MLMPDATCVNSRTCAMINCQYSCEDTEEGPQCLCPSSGLRLAPNGRDCLDIDECASGKVICP
YNRRCVNTFGSYYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFNTQGSFKCKCK
QGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKKLLAHKNSMKKKAKIKNVTPEPTRTPTPK
VNLQPFNYEEIVSRGGNSHGGKKGNEEK

Signal peptide:

amino acids 1-21

EGF-like domain cysteine pattern signature.

amino acids 80-91

Calcium-binding EGF-like domains

amino acids 103-124, 230-251 and 185-206

FIGURE 46

GGGAGCTGCTGCTGTGGCTGCTGGTGCTGTGCGCGCTGCTCCTGCTCTTGGTGCAGCTGCTG

CGCTTCCTGAGGGCTGACGGCGACCTGACGCTACTATGGGCCGAGTGGCAGGGACGACGCCC

AGAATGGGAGCTGACTGAT**ATG**GTGGTGTGGGTGACTGGAGCCTCGAGTGGAATTGGTGAGG

AGCTGGCTTACCAGTTGTCTAAACTAGGAGTTTCTCTTGTGCTGTCAGCCAGAAGAGTGCAT

GAGCTGGAAAGGGTGAAAAGAAGATGCCTAGAGAATGGCAATTTAAAAGAAAAAGATATACT

TGTTTTGCCCCTTGACCTGACCGACACTGGTTCCCATGAAGCGGCTACCAAAGCTGTTCTCC

AGGAGTTTGGTAGAATCGACATTCTGGTCAACAATGGTGGAATGTCCCAGCGTTCTCTGTGC

ATGGATACCAGCTTGGATGTCTACAGAAAGCTAATAGAGCTTAACTACTTAGGGACGGTGTC

CTTGACAAAATGTGTTCTGCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA

ATAGCATCCTGGGTATCATATCTGTACCTCTTTCCATTGGATACTGTGCTAGCAAGCATGCT

CTCCGGGGTTTTTTTAATGGCCTTCGAACAGAACTTGCCACATACCCAGGTATAATAGTTTC

TAACATTTGCCCAGGACCTGTGCAATCAAATATTGTGGAGAATTCCCTAGCTGGAGAAGTCA

CAAAGACTATAGGCAATAATGGAGACCAGTCCCACAAGATGACAACCAGTCGTTGTGTGCGG

CTGATGTTAATCAGCATGGCCAATGATTTGAAAGAAGTTTGGATCTCAGAACAACCTTTCTT

GTTAGTAACATATTTGTGGCAATACATGCCAACCTGGGCCTGGTGGATAACCAACAAGATGG

GGAAGAAAAGGATTGAGAACTTTAAGAGTGGTGTGGATGCAGACTCTTCTTATTTTAAAATC

TTTAAGACAAAACA**TGA**CTGAAAAGAGCACCTGTACTTTTCAAGCCACTGGAGGGAGAAATG

GAAAACATGAAAACAGCAATCTTCTTATGCTTCTGAATAATCAAAGACTAATTTGTGATTTT

ACTTTTTAATAGATATGACTTTGCTTCCAACATGGAATGAAATAAAAAATAAATAATAAAAG

ATTGCCATGAATCTTGCAAAA

FIGURE 47

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343

><subunit 1 of 1, 289 aa, 1 stop

><MW: 32268, pI: 9.21, NX(S/T): 0

MVVWVTGASSGIGEELAYQLSKLGVSLVLSARRVHELERVKRRCLENGNLKEKDILVLPLDL

TDTGSHEAATKAVLQEFGRIDILVNNGGMSQRSLCMDTSLDVYRKLIELNYLGTVSLTKCVL

PHMIERKQGKIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIVSNICPGP

VQSNIVENSLAGEVTKTIGNNGDQSHKMTTSRCVRLMLISMANDLKEVWISEQPFLLVTYLW

QYMPTWAWWITNKMGKKRIENFKSGVDADSSYFKIFKTKHD

Important Features:

Signal Peptide:

amino acids 1-31

Transmembrane domain:

amino acids 136-157

Tyrosine kinase phosphorylation site.

106-113 and 107-114

Homologous region to Short-chain alcohol dehydrogenase

amino acids 80-90, 131-168, 1-13 and 176-185

FIGURE 48

GCGACGTGGGCACCGCCATCAGCTGTTCGCGCGTCTTCTCCTCCAGGTGGGGCAGGGGTTTC
GGGCTGGTGGAGCATGTGCTGGGACAGGACAGCATCCTCAATCAATCCAACAGCATATTCGG
TTGCATCTTCTACACACTACAGCTATTGTTAGGTTGCCTGCGGACACGCTGGGCCTCTGTCC
TG**ATG**CTGCTGAGCTCCCTGGTGTCTCTCGCTGGTTCTGTCTACCTGGCCTGGATCCTGTTC
TTCGTGCTCTATGATTTCTGCATTGTTTGTATCACCACCTATGCTATCAACGTGAGCCTGAT
GTGGCTCAGTTTCCGGAAGGTCCAAGAACCCCAGGGCAAGGCTAAGAGGCACTGAGCCCTCA
ACCCAAGCCAGGCTGACCTCATCTGCTTTGCTTTGGTCTTCAAGCCGCTCAGCGTGCCTGTG
GACAGCGTGGCCCCGGCCCCCCCAAGCCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGCCC
TGGCAGGCCAGTGTGAGGAGGCAAGGAGCCCACATCTGCAGCGGCTCCCTGGTGGCAGACAC
CTGGGTCCTCACTGCTGCCCACTGCTTTGAAAAGGCAGCAGCAACAGAACTGAATTCCTGGT
CAGTGGTCCTGGGTTCTCTGCAGCGTGAGGGACTCAGCCCTGGGGCCGAAGAGGTGGGGGTG
GCTGCCCTGCAGTTGCCCAGGGCCTATAACCACTACAGCCAGGGCTCAGACCTGGCCCTGCT
GCAGCTCGCCCACCCCACGACCCACACACCCCTCTGCCTGCCCCAGCCCGCCCATCGCTTCC
CCTTTGGAGCCTCCTGCTGGGCCACTGGCTGGGATCAGGACACCAGTGATGCTCCTGGGACC
CTACGCAATCTGCGCCTGCGTCTCATCAGTCGCCCCACATGTAACTGTATCTACAACCAGCT
GCACCAGCGACACCTGTCCAACCCGGCCCGGCCTGGGATGCTATGTGGGGGCCCCCAGCCTG
GGGTGCAGGGCCCCTGTCAGGGAGATTCCGGGGGCCCTGTGCTGTGCCTCGAGCCTGACGGA
CACTGGGTTCAGGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCCAGGAGGACGCTCCTGT
GCTGCTGACCAACACAGCTGCTCACAGTTCCTGGCTGCAGGCTCGAGTTCAGGGGGCAGCTT
TCCTGGCCCAGAGCCCAGAGACCCCGGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGA
TCCTTGAGGACAGCAGGTCCCCAGGCAGGAGCACCCTCCCCATGGCCCTGGGAGGCCAGGCT
GATGCACCAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTCAGAGGAGGCGGTGCTAACTG
CTGCCCACTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGACCAGA
CCGGAGGAGTGGGGCCTGAAGCAGCTCATCCTGCATGGAGCCTACACCCACCCTGAGGGGGG
CTACGACATGGCCCTCCTGCTGCTGGCCCAGCCTGTGACACTGGGAGCCAGCCTGCGGCCCC
TCTGCCTGCCCTATCCTGACCACCACCTGCCTGATGGGGAGCGTGGCTGGGTTCTGGGACGG
GCCCGCCCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCGTGACCCTCCTGGGGCCTAG
GGCCTGCAGCCGGCTGCATGCAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGGATGG
TGTGTACCAGTGCTGTGGGTGAGCTGCCCAGCTGTGAGGGCCTGTCTGGGGCACCACTGGTG
CATGAGGTGAGGGGCACATGGTTCCTGGCCGGGCTGCACAGCTTCGGAGATGCTTGCCAAGG
CCCCGCCAGGCCGGCGGTCTTCACCGCGCTCCCTGCCTATGAGGACTGGGTCAGCAGTTTGG
ACTGGCAGGTCTACTTCGCCGAGGAACCAGAGCCCGAGGCTGAGCCTGGAAGCTGCCTGGCC
AACATAAGCCAACCAACCAGCTGC**TGA**CAGGGGACCTGGCCATTCTCAGGACAAGAGAATGC
AGGCAGGCAAATGGCATTACTGCCCCTGTCCTCCCCACCCTGTCATGTGTGATTCCAGGCAC
CAGGGCAGGCCCAGAAGCCCAGCAGCTGTGGGAAGGAACCTGCCTGGGGCCACAGGTGCCCA
CTCCCCACCCTGCAGGACAGGGGTGTCTGTGGACACTCCCACACCCAACTCTGCTACCAAGC
AGGCGTCTCAGCTTTCCTCCTCCTTTACTCTTTCAGATACAATCACGCCAGCCACGTTGTTT
TGAAAATTTCTTTTTTTGGGGGGCAGCAGTTTTCCTTTTTTTAAACTTAAATAAATTGTTAC
AAAATAAAA

FIGURE 49

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571

MLLSSLVSLAGSVYLAWILFFVLYDFCIVCITTYAINVSLMWLSFRKVQEPQGKAKRHGNTV
PGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAAATELNSWSVVLGSLQREGLSPGA
EEVGVAALQLPRAYNHYSQGSDLALLQLAHPTTHTPLCLPQPAHRFPFGASCWATGWDQDTS
DAPGTLRNLRLRLISRPTCNCIYNQLHQRHLSNPARPGMLCGGPQPGVQGPCQGDSGGPVLC
LEPDGHWVQAGIISFASSCAQEDAPVLLTNTAAHSSWLQARVQGAAFLAQSPETPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQGQLACGGALVSEEAVLTAAHCFIGRQAPEEWSV
GLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYPDHHLPDGERG
WVLGRARPGAGISSLQTVPVTLLGPRACSRLHAAPGGDGSPILPGMVCTSAVGELPSCEGLS
GAPLVHEVRGTWFLAGLHSFGDACQGPARPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAEP
GSCLANISQPTSC

Important features:
Signal peptide:
amino acids 1-15

Homologous region to Serine proteases, trypsin family
amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.
amino acids 37-40 and 564-567

Kringle domains
amino acids 79-96, 343-360 and 235-247

FIGURE 50

CGGGCCGCCCCCGGCCCCCATTCGGGCCGGGCCTCGCTGCGGCGGCGACTGAGCCAGGCTGG

GCCGCGTCCCTGAGTCCCAGAGTCGGCGCGGCGCGGCAGGGGCAGCCTTCCACCACGGGGAG

CCCAGCTGTCAGCCGCCTCACAGGAAG**ATG**CTGCGTCGGCGGGGCAGCCCTGGCATGGGTGT

GCATGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG

TCCCTGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCCTGTGCTGCTCCTTCTCC

CCTGAGCCTGGCTTCAGCCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATACCAAACA

GCTGGTGCACAGCTTTGCTGAGGGCCAGGACCAGGGCAGCGCCTATGCCAACCGCACGGCCC

TCTTCCCGGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGCGTGCGTGTGGCG

GACGAGGGCAGCTTCACCTGCTTCGTGAGCATCCGGGATTTCGGCAGCGCTGCCGTCAGCCT

GCAGGTGGCCGCTCCCTACTCGAAGCCCAGCATGACCCTGGAGCCCAACAAGGACCTGCGGC

CAGGGGACACGGTGACCATCACGTGCTCCAGCTACCAGGGCTACCCTGAGGCTGAGGTGTTC

TGGCAGGATGGGCAGGGTGTGCCCCTGACTGGCAACGTGACCACGTCGCAGATGGCCAACGA

GCAGGGCTTGTTTGATGTGCACAGCGTCCTGCGGGTGGTGCTGGGTGCGAATGGCACCTACA

GCTGCCTGGTGCGCAACCCCGTGCTGCAGCAGGATGCGCACRGCTCTGTCACCATCACAGGG

CAGCCTATGACATTCCCCCCAGAGGCCCTGTGGGTGACCGTGGGGCTGTCTGTCTGTCTCAT

TGCACTGCTGGTGGCCCTGGCTTTCGTGTGCTGGAGAAAGATCAAACAGAGCTGTGAGGAGG

AGAATGCAGGAGCTGAGGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCT

CTGAAACACTCTGACAGCAAAGAAGATGATGGACAAGAAATAGCC**TGA**CCATGAGGACCAGG

GAGCTGCTACCCCTCCCTACAGCTCCTACCCTCTGGCTGCAATGGGGCTGCACTGTGAGCCC

TGCCCCCAACAGATGCATCCTGCTCTGACAGGTGGGCTCCTTCTCCAAAGGATGCGATACAC

AGACCACTGTGCAGCCTTATTTCTCCAATGGACATGATTCCCAAGTCATCCTGCTGCCTTTT

TTCTTATAGACACAATGAACAGACCACCCACAACCTTAGTTCTCTAAGTCATCCTGCCTGCT

GCCTTATTTCACAGTACATACATTTCTTAGGGACACAGTACACTGACCACATCACCACCCTC

TTCTTCCAGTGCTGCGTGGACCATCTGGCTGCCTTTTTTCTCCAAAAGATGCAATATTCAGA

CTGACTGACCCCCTGCCTTATTTCACCAAAGACACGATGCATAGTCACCCCGGCCTTGTTTC

TCCAATGGCCGTGATACACTAGTGATCATGTTCAGCCCTGCTTCCACCTGCATAGAATCTTT

TCTTCTCAGACAGGGACAGTGCGGCCTCAACATCTCCTGGAGTCTAGAAGCTGTTTCCTTTC

CCCTCCTTCCTCCCTGCCCCAAGTGAAGACAGGGCAGGGCCAGGAATGCTTTGGGGACACCG

AGGGGACTGCCCCCCACCCCCACCATGGTGCTATTCTGGGGCTGGGGCAGTCTTTTCCTGGC

TTGCCTCTGGCCAGCTCCTGGCCTCTGGTAGAGTGAGACTTCAGACGTTCTGATGCCTTCCG

GATGTCATCTCTCCCTGCCCCAGGAATGGAAGATGTGAGGACTTCTAATTTAAATGTGGGAC

TCGGAGGGATTTTGTAAACTGGGGGTATATTTTGGGGAAAATAAATGTCTTTGTAAAAAAAA

AAAAAAAAAAAAA

FIGURE 51

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386

><subunit 1 of 1, 316 aa, 1 stop, 1 unknown

><MW: -1, pI: 4.62, NX(S/T): 4

MLRRRGSPGMGVHVGAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPEPGFSLAQ

LNLIWQLTDTKQLVHSFAEGQDQGSAYANRTALFPDLLAQGNASLRLQRVRVADEGSFTCFV

SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDTVTITCSSYQGYPEAEVFWQDGQGVPL

TGNVTTSQMANEQGLFDVHSVLRVVLGANGTYSCLVRNPVLQQDAHXSVTITGQPMTFPPEA

LWVTVGLSVCLIALLVALAFVCWRKIKQSCEEENAGAEDQDGEGEGSKTALQPLKHSDSKED

DGQEIA

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 251-270

N-glycosylation site.

amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC

amino acids 217-234

FIGURE 52

TTCGTGACCCTTGAGAAAAGAGTTGGTGGTAAATGTGCCACGTCTTCTAAGAAGGGGGAGTC
CTGAACTTGTCTGAAGCCCTTGTCCGTAAGCCTTGAACTACGTTCTTAAATCTATGAAGTCG
AGGGACCTTTCGCTGCTTTTGTAGGGACTTCTTTCCTTGCTTCAGCAAC**ATG**AGGCTTTTCT
TGTGGAACGCGGTCTTGACTCTGTTCGTCACTTCTTTGATTGGGGCTTTGATCCCTGAACCA
GAAGTGAAAATTGAAGTTCTCCAGAAGCCATTCATCTGCCATCGCAAGACCAAAGGAGGGGA
TTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTCACTCCACTC
ACAAACATAACAATGGTCAGCCCATTTGGTTTACCCTGGGCATCCTGGAGGCTCTCAAAGGT
TGGGACCAGGGCTTGAAAGGAATGTGTGTAGGAGAGAAGAGAAAGCTCATCATTCCTCCTGC
TCTGGGCTATGGAAAAGAAGGAAAAGGTAAAATTCCCCCAGAAAGTACACTGATATTTAATA
TTGATCTCCTGGAGATTCGAAATGGACCAAGATCCCATGAATCATTCCAAGAAATGGATCTT
AATGATGACTGGAAACTCTCTAAAGATGAGGTTAAAGCATATTTAAAGAAGGAGTTTGAAAA
ACATGGTGCGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTTGATAAAG
AAGATGAAGACAAAGATGGGTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTA
TAGAGATACATCTACCCTTTTAATATAGCACTCATCTTTCAAGAGAGGGCAGTCATCTTTAA
AGAACATTTTATTTTTATACAATGTTCTTTCTTGCTTTGTTTTTTATTTTTATATATTTTTT
CTGACTCCTATTTAAAGAACCCCTTAGGTTTCTAAGTACCCATTTCTTTCTGATAAGTTATT
GGGAAGAAAAAGCTAATTGGTCTTTGAATAGAAGACTTCTGGACAATTTTTCACTTTCACAG
ATATGAAGCTTTGTTTTACTTTCTCACTTATAAATTTAAAATGTTGCAACTGGGAATATACC
ACGACATGAGACCAGGTTATAGCACAAATTAGCACCCTATATTTCTGCTTCCCTCTATTTTC
TCCAAGTTAGAGGTCAACATTTGAAAAGCCTTTTGCAATAGCCCAAGGCTTGCTATTTTCAT
GTTATAATGAAATAGTTTATGTGTAACTGGCTCTGAGTCTCTGCTTGAGGACCAGAGGAAAA
TGGTTGTTGGACCTGACTTGTTAATGGCTACTGCTTTACTAAGGAGATGTGCAATGCTGAAG
TTAGAAACAAGGTTAATAGCCAGGCATGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAG
GCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTCGAGACCAGCCTGACCAACACGGAGAAA
CCCTATCTCTACTAAAAATACAAAGTAGCCCGGCGTGGTGATGCGTGCCTGTAATCCCAGCT
ACCCAGGAAGGCTGAGGCGGCAGAATCACTTGAACCCGAGGCCGAGGTTGCGGTAAGCCGAG
ATCACCTNCAGCCTGGACACTCTGTCTCGAAAAAAGAAAAGAACACGGTTAATACCATATNA
ATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGGCTCCTAGTGAT
TGGTGGCCTATTATGATAAATAGGACAAATCATTTATGTGTGAGTTTCTTTGTAATAAAATG
TATCAATATGTTATAGATGAGGTAGAAAGTTATATTTATATTCAATATTTACTTCTTAAGGC
TAGCGGAATATCCTTCCTGGTTCTTTAATGGGTAGTCTATAGTATATTATACTACAATAACA
TTGTATCATAAGATAAAGTAGTAAACCAGTCTACATTTTCCCATTTCTGTCTCATCAAAAAC
TGAAGTTAGCTGGGTGTGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGGGCCAAGGAGGG
TGGATCACTTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTCTA
CTAAAAATACAAAAATTAGCCAGGCGTGGTGGTGCACACCTGTAGTCCCAGCTACTCGGGAG
GCTGAGACAGGAGATTTGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCC
ACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAGAAGCAGA
CCTACAGCAGCTACTATTGAATAAATACCTATCCTGGATTTT

FIGURE 53

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44194
><subunit 1 of 1, 211 aa, 1 stop
><MW: 24172, pI: 5.99, NX(S/T): 1

MRLFLWNAVLTLFVTSLIGALIPEPEVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKDGSL

FHSTHKHNNGQPIWFTLGILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPEST

LIFNIDLLEIRNGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKHGAVVNESHHDALVED

IFDKEDEDKDGFISAREFTYKHDEL

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 176-179

Casein kinase II phosphorylation site.

amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.

amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 78-114 and 118-131

EF-hand calcium-binding domain.

amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain

amino acids 183-203

FIGURE 54

AATAAAGCTTCCTTAATGTTGTATATGTCTTTGAAGTACATCCGTGCATTTTTTTTTAGCAT
CCAACCATTCCTCCCTTGTAGTTCTCGCCCCCTCAAATCACCCTCTCCCGTAGCCCACCCGA
CTAACATCTCAGTCTCTGAAA**ATG**CACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCT
CACGGGGCTCAGTCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCACAGTAC
CTGCCACCCTCAACGTCCTCAATGGCTCTGACGCCCGCCTGCCCTGCACCTTCAACTCCTGC
TACACAGTGAACCACAAACAGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACTGCTC
TGAGGAGATGTTCCTCCAGTTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAG
ACCGCGTGGAGTTCTCAGGGAACCCCAGCAAGTACGATGTGTCGGTGATGCTGAGAAACGTG
CAGCCGGAGGATGAGGGGATTTACAACTGCTACATCATGAACCCCCCTGACCGCCACCGTGG
CCATGGCAAGATCCATCTGCAGGTCCTCATGGAAGAGCCCCCTGAGCGGGACTCCACGGTGG
CCGTGATTGTGGGTGCCTCCGTCGGGGGCTTCCTGGCTGTGGTCATCTTGGTGCTGATGGTG
GTCAAGTGTGTGAGGAGAAAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACCGAGGA
GGAGGGCAAGACGGACGGTGAAGGCAACCCGGATGATGGCGCCAAG**TAG**TGGGTGGCCGGCC
CTGCAGCCTCCCGTGTCCCGTCTCCTCCCCTCTCCGCCCTGTACAGTGACCCTGCCTGCTCG
CTCTTGGTGTGCTTCCCGTGACCTAGGACCCCAGGGCCCACCTGGGGCCTCCTGAACCCCCG
ACTTCGTATCTCCCACCCTGCACCAAGAGTGACCCACTCTCTTCCATCCGAGAAACCTGCCA
TGCTCTGGGACGTGTGGGCCCTGGGGAGAGGAGAGAAAGGGCTCCCACCTGCCAGTCCCTGG
GGGGAGGCAGGAGGCACATGTGAGGGTCCCCAGAGAGAAGGGAGTGGGTGGGCAGGGGTAGA
GGAGGGGCCGCTGTCACCTGCCCAGTGCTTGCCTGGCAGTGGCTTCAGAGAGGACCTGGTGG
GGAGGGAGGGCTTTCCTGTGCTGACAGCGCTCCCTCAGGAGGGCCTTGGCCTGGCACGGCTG
TGCTCCTCCCCTGCTCCCAGCCCAGAGCAGCCATCAGGCTGGAGGTGACGATGAGTTCCTGA
AACTTGGAGGGGCATGTTAAAGGGATGACTGTGCATTCCAGGGCACTGACGGAAAGCCAGGG
CTGCAGGCAAAGCTGGACATGTGCCCTGGCCCAGGAGGCCATGTTGGGCCCTCGTTTCCATT
GCTAGTGGCCTCCTTGGGGCTCCTGTTGGCTCCTAATCCCTTAGGACTGTGGATGAGGCCAG
ACTGGAAGAGCAGCTCCAGGTAGGGGGCCATGTTTCCCAGCGGGGACCCACCAACAGAGGCC
AGTTTCAAAGTCAGCTGAGGGGCTGAGGGGTGGGGCTCCATGGTGAATGCAGGTTGCTGCAG
GCTCTGCCTTCTCCATGGGGTAACCACCCTCGCCTGGGCAGGGGCAGCCAAGGCTGGGAAAT
GAGGAGGCCATGCACAGGGTGGGGCAGCTTTCTTTGGGGCTTCAGTGAGAACTCTCCCAGTT
GCCCTTGGTGGGGTTTCCACCTGGCTTTTGGCTACAGAGAGGGAAGGGAAAGCCTGAGGCCG
GCATAAGGGGAGGCCTTGGAACCTGAGCTGCCAATGCCAGCCCTGTCCCATCTGCGGCCACG
CTACTCGCTCCTCTCCCAACAACTCCCTTCGTGGGGACAAAAGTGACAATTGTAGGCCAGGC
ACAGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGTGGATTACCTCCAT
CTGTTTAGTAGAAATGGGCAAAACCCCATCTCTACTAAAAATACAAGAATTAGCTGGGCGTG
GTGGCGTGTGCCTGTAATCCCAGCTATTTGGGAGGCTGAGGCAGGAGAATCGCTTGAGCCCG
GGAAGCAGAGGTTGCAGTGAACTGAGATAGTGATAGTGCCACTGCAATTCAGCCTGGGTGAC
ATAGAGAGACTCCATCTCAAAAAAAA

FIGURE 55

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415

<subunit 1 of 1, 215 aa, 1 stop

<MW: 24326, pI: 6.32, NX(S/T): 4

MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNVLNGSDARLPCTFNSCYTVNHKQ

FSLNWTYQECNNCSEEMFLQFRMKIINLKLERFQDRVEFSGNPSKYDVSVMLRNVQPEDEGI

YNCYIMNPPDRHRGHGKIHLQVLMEEPPERDSTVAVIVGASVGGFLAVVILVLMVVKCVRRK

KEQKLSTDDLKTEEEGKTDGEGNPDDGAK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 161-179

Immunoglobulin-like fold:

amino acids 83-127

N-glycosylation sites.

amino acids 42-45, 66-69 and 74-77

FIGURE 56

GTTGTATATGTCCTGAAGTACATCCGTGCATTTTTTTTAGCATCCAACCATCCTCCCTTGTA

GTTCTCGCCCCCTCAAATCACCTTCTCCCTTAGCCCACCCNACTAACATCTCAGTCTCTGAA

AATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCTCACGGGGCTCAGTCTCTTTT

TCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCCACAGTACCTGNCCACCCTCAACGTCC

TCAATGGCTCTGACGCCCGCCTGCCCTGCCCTTCAACTCCTGCTACACAGTGAACCACAAAC

AGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACTGCTCTGAGGAGATGTTCCTCCAG

TTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAGACCGCGTGGAGTTCTCAGG

GAACCCCAGCAAGTACGATGTGTCGGTGATGCTGAGAAACGTGCAGCCGGAGGATGAGGGGA

TTTACAACTGCTACATCATGAACCCCCC

FIGURE 57

TCACGGGGCTCATCTCTTTTTTCTCTTTGGTGCCCACCAGGACGGAGCATGGAGGTNCACATA

CCTGCCACCCTCAACGTCCTCAATGGCTTTGACGCCCGCCTGCCCTGCACCTTCAACTCCNG

CTACACAGTGAACCACAAACAGTTCTCCCTGAACTGGATTTACCAGGAGTGCAACAACTGGC

TCTGAGGAGATGTTCCTCCAGTTCCCGCATGGAAGATCATTTAACCTGAAAGCTGGAAGCGG

TTTTCAAGAACCGCGTGGAAGTTTCTCAGGGAACCCCAGCAAGTACGATGTGTCGGTGATGC

TGAGAAACGTGCAGCCGGAGGATGAGGGGATTTACAACTGCTACATCATGAACCCCCC

FIGURE 58

```
TGCGGCGACCGTCGTACACCATGGGCCTCCACCTCCGCCCCTACCGTGTGGGGCTGCTCCCGGATGGCCTCCTGT
TCCTCTTGCTGCTGCTAATGCTGCTCGCGGACCCAGCGCTCCCGGCCGGACGTCACCCCCCAGTGGTGCTGGTCC
CTGGTGATTTGGGTAACCAACTGGAAGCCAAGCTGGACAAGCCGACAGTGGTGCACTACCTCTGCTCCAAGAAGA
CCGAAAGCTACTTCACAATCTGGCTGAACCTGGAACTGCTGCTGCCTGTCATCATTGACTGCTGGATTGACAATA
TCAGGCTGGTTTACAACAAAACATCCAGGGCCACCCAGTTTCCTGATGGTGTGGATGTACGTGTCCCTGGCTTTG
GGAAGACCTTCTCACTGGAGTTCCTGGACCCCAGCAAAAGCAGCGTGGGTTCCTATTTCCACACCATGGTGGAGA
GCCTTGTGGGCTGGGGCTACACACGGGGTGAGGATGTCCGAGGGGCTCCCTATGACTGGCGCCGAGCCCCAAATG
AAAACGGGCCCTACTTCCTGGCCCTCCGCGAGATGATCGAGGAGATGTACCAGCTGTATGGGGGCCCCGTGGTGC
TGGTTGCCCACAGTATGGGCAACATGTACACGCTCTACTTTCTGCAGCGGCAGCCGCAGGCCTGGAAGGACAAGT
ATATCCGGGCCTTCGTGTCACTGGGTGCGCCCTGGGGGGGCGTGGCCAAGACCCTGCGCGTCCTGGCTTCAGGAG
ACAACAACCGGATCCCAGTCATCGGGCCCCTGAAGATCCGGGAGCAGCAGCGGTCAGCTGTCTCCACCAGCTGGC
TGCTGCCCTACAACTACACATGGTCACCTGAGAAGGTGTTCGTGCAGACACCCACAATCAACTACACACTGCGGG
ACTACCGCAAGTTCTTCCAGGACATCGGCTTTGAAGATGGCTGGCTCATGCGGCAGGACACAGAAGGGCTGGTGG
AAGCCACGATGCCACCTGGCGTGCAGCTGCACTGCCTCTATGGTACTGGCGTCCCCACACCAGACTCCTTCTACT
ATGAGAGCTTCCCTGACCGTGACCCTAAAATCTGCTTTGGTGACGGCGATGGTACTGTGAACTTGAAGAGTGCCC
TGCAGTGCCAGGCCTGGCAGAGCCGCCAGGAGCACCAAGTGTTGCTGCAGGAGCTGCCAGGCAGCGAGCACATCG
AGATGCTGGCCAACGCCACCACCCTGGCCTATCTGAAACGTGTGCTCCTTGGGCCCTGACTCCTGTGCCACAGGA
CTCCTGTGGCTCGGCCGTGGACCTGCTGTTGGCCTCTGGGGCTGTCATGGCCCACGCGTTTTGCAAAGTTTGTGA
CTCACCATTCAAGGCCCCGAGTCTTGGACTGTGAAGCATCTGCCATGGGGAAGTGCTGTTTGTTATCCTTTCTCT
GTGGCAGTGAAGAAGGAAGAAATGAGAGTCTAGACTCAAGGGACACTGGATGGCAAGAATGCTGCTGATGGTGGA
ACTGCTGTGACCTTAGGACTGGCTCCACAGGGTGGACTGGCTGGGCCCTGGTCCCAGTCCCTGCCTGGGGCCATG
TGTCCCCCTATTCCTGTGGGCTTTTCATACTTGCCTACTGGGCCCTGGCCCCGCAGCCTTCCTATGAGGGATGTT
ACTGGGCTGTGGTCCTGTACCCAGAGGTCCCAGGGATCGGCTCCTGGCCCCTCGGGTGACCCTTCCCACACACCA
GCCACAGATAGGCCTGCCACTGGTCATGGGTAGCTAGAGCTGCTGGCTTCCCTGTGGCTTAGCTGGTGGCCAGCC
TGACTGGCTTCCTGGGCGAGCCTAGTAGCTCCTGCAGGCAGGGGCAGTTTGTTGCGTTCTTCGTGGTTCCCAGGC
CCTGGGACATCTCACTCCACTCCTACCTCCCTTACCACCAGGAGCATTCAAGCTCTGGATTGGGCAGCAGATGTG
CCCCCAGTCCCGCAGGCTGTGTTCCAGGGGCCCTGATTTCCTCGGATGTGCTATTGGCCCCAGGACTGAAGCTGC
CTCCCTTCACCCTGGGACTGTGGTTCCAAGGATGAGAGCAGGGGTTGGAGCCATGGCCTTCTGGGAACCTATGGA
GAAAGGGAATCCAAGGAAGCAGCCAAGGCTGCTCGCAGCTTCCCTGAGCTGCACCTCTTGCTAACCCCACCATCA
CACTGCCACCCTGCCCTAGGGTCTCACTAGTACCAAGTGGGTCAGCACAGGGCTGAGGATGGGGCTCCTATCCAC
CCTGGCCAGCACCCAGCTTAGTGCTGGGACTAGCCCAGAAACTTGAATGGGACCCTGAGAGAGCCAGGGGTCCCC
TGAGGCCCCCCTAGGGGCTTTCTGTCTGCCCCAGGGTGCTCCATGGATCTCCCTGTGGCAGCAGGCATGGAGAGT
CAGGGCTGCCTTCATGGCAGTAGGCTCTAAGTGGGTGACTGGCCACAGGCCGAGAAAAGGGTACAGCCTCTAGGT
GGGGTTCCCAAAGACGCCTTCAGGCTGGACTGAGCTGCTCTCCCACAGGGTTTCTGTGCAGCTGGATTTTCTCTG
TTGCATACATGCCTGGCATCTGTCTCCCCTTGTTCCTGAGTGGCCCCACATGGGGCTCTGAGCAGGCTGTATCTG
GATTCTGGCAATAAAAGTACTCTGGATGCTGTAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 59

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44189

><subunit 1 of 1, 412 aa, 1 stop

><MW: 46658, pI: 6.65, NX(S/T): 4

MGLHLRPYRVGLLPDGLLFLLLLLMLLADPALPAGRHPPVVLVPGDLGNQLEAKLDKPTVVH
YLCSKKTESYFTIWLNLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGVDVRVPGFGKTFSL
EFLDPSKSSVGSYFHTMVESLVGWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEEMYQ
LYGGPVVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNNRI
PVIGPLKIREQQRSAVSTSWLLPYNYTWSPEKVFVQTPTINYTLRDYRKFFQDIGFEDGWLM
RQDTEGLVEATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGDGTVNLKSALQCQ
AWQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLLGP

Important features:

Signal peptide:

amino acids 1-28

Potential lipid substrate binding site:

amino acids 147-164

N-glycosylation sites.

amino acids 99-102, 273-276, 289-292 and 398-401

Lipases, serine proteins

amino acids 189-201

Beta-transducin family Trp-Asp repeat

amino acids 353-365

FIGURE 60

CGGACGCGTGGGCGGACGCGTGGGGCGGCGGCAGCGGCGGCGACGGCGAC**ATG**GAGAGCGGG

GCCTACGGCGCGGCCAAGGCGGGCGGCTCCTTCGACCTGCGGCGCTTCCTGACGCAGCCGCA

GGTGGTGGCGCGCGCCGTGTGCTTGGTCTTCGCCTTGATCGTGTTCTCCTGCATCTATGGTG

AGGGCTACAGCAATGCCCACGAGTCTAAGCAGATGTACTGCGTGTTCAACCGCAACGAGGAT

GCCTGCCGCTATGGCAGTGCCATCGGGGTGCTGGCCTTCCTGGCCTCGGCCTTCTTCTTGGT

GGTCGACGCGTATTTCCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTCATTGGTG

ACCTGCTCTTCTCAGCTCTCTGGACCTTCCTGTGGTTTGTTGGTTTCTGCTTCCTCACCAAC

CAGTGGGCAGTCACCAACCCGAAGGACGTGCTGGTGGGGGCCGACTCTGTGAGGGCAGCCAT

CACCTTCAGCTTCTTTTCCATCTTCTCCTGGGGTGTGCTGGCCTCCCTGGCCTACCAGCGCT

ACAAGGCTGGCGTGGACGACTTCATCCAGAATTACGTTGACCCCACTCCGGACCCCAACACT

GCCTACGCCTCCTACCCAGGTGCATCTGTGGACAACTACCAACAGCCACCCTTCACCCAGAA

CGCGGAGACCACCGAGGGCTACCAGCCGCCCCCTGTGTAC**TGA**GTGGCGGTTAGCGTGGGAA

GGGGGACAGAGAGGGCCCTCCCCTCTGCCCTGGACTTTCCCATCAGCCTCCTGGAACTGCCA

GCCCCTCTCTTTCACCTGTTCCATCCTGTGCAGCTGACACACAGCTAAGGAGCCTCATAGCC

TGGCGGGGGCTGGCAGAGCCACACCCCAAGTGCCTGTGCCCAGAGGGCTTCAGTCAGCCGCT

CACTCCTCCAGGGCACTTTTAGGAAAGGGTTTTTAGCTAGTGTTTTTCCTCGCTTTTAATGA

CCTCAGCCCCGCCTGCAGTGGCTAGAAGCCAGCAGGTGCCCATGTGCTACTGACAAGTGCCT

CAGCTTCCCCCCGGCCCGGGTCAGGCCGTGGGAGCCGCTATTATCTGCGTTCTCTGCCAAAG

ACTCGTGGGGGCCATCACACCTGCCCTGTGCAGCGGAGCCGGACCAGGCTCTTGTGTCCTCA

CTCAGGTTTGCTTCCCCTGTGCCCACTGCTGTATGATCTGGGGGCCACCACCCTGTGCCGGT

GGCCTCTGGGCTGCCTCCCGTGGTGTGAGGGCGGGGCTGGTGCTCATGGCACTTCCTCCTTG

CTCCCACCCCTGGCAGCAGGGAAGGGCTTTGCCTGACAACACCCAGCTTTATGTAAATATTC

TGCAGTTGTTACTTAGGAAGCCTGGGGAGGGCAGGGGTGCCCCATGGCTCCCAGACTCTGTC

TGTGCCGAGTGTATTATAAAATCGTGGGGGAGATGCCCGGCCTGGGATGCTGTTTGGAGACG

GAATAAATGTTTTCTCATTCAAAG

FIGURE 61

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48304

<subunit 1 of 1, 224 aa, 1 stop

<MW: 24810, pI: 4.75, NX(S/T): 1

MESGAYGAAKAGGSFDLRRFLTQPQVVARAVCLVFALIVFSCIYGEGYSNAHESKQMYCVFN

RNEDACRYGSAIGVLAFLASAFFLVVDAYFPQISNATDRKYLVIGDLLFSALWTFLWFVGFC

FLTNQWAVTNPKDVLVGADSVRAAITFSFFSIFSWGVLASLAYQRYKAGVDDFIQNYVDPTP

DPNTAYASYPGASVDNYQQPPFTQNAETTEGYQPPPVY

Important features:

Type II Transmembrane domain:

amino acids 1-45

Other transmembrane domains:

amino acids 74-90, 108-126 and 145-161

N-glycosylation site.

amino acids 97-100

FIGURE 62

GAGCCACCTACCCTGCTCCGAGGCCAGGCCTGCAGGGCCTCATCGGCCAGAGGGTGATCAGTGAGCAGAAGG**ATG**
CCCGTGGCCGAGGCCCCCCAGGTGGCTGGCGGGCAGGGGGACGGAGGTGATGGCGAGGAAGCGGAGCCAGAGGGG
ATGTTCAAGGCCTGTGAGGACTCCAAGAGAAAAGCCCGGGGCTACCTCCGCCTGGTGCCCCTGTTTGTGCTGCTG
GCCCTGCTCGTGCTGGCTTCGGCGGGGGTGCTACTCTGGTATTTCCTAGGGTACAAGGCGGAGGTGATGGTCAGC
CAGGTGTACTCAGGCAGTCTGCGTGTACTCAATCGCCACTTCTCCCAGGATCTTACCCGCCGGGAATCTAGTGCC
TTCCGCAGTGAAACCGCCAAAGCCCAGAAGATGCTCAAGGAGCTCATCACCAGCACCCGCCTGGGAACTTACTAC
AACTCCAGCTCCGTCTATTCCTTTGGGGAGGGACCCCTCACCTGCTTCTTCTGGTTCATTCTCCAAATCCCCGAG
CACCGCCGGCTGATGCTGAGCCCCGAGGTGGTGCACTGCTGGTGGAGGAGCTGCTGTCCACAGTCAACAGC
TCGGCTGCCGTCCCCTACAGGGCCGAGTACGAAGTGGACCCCGAGGGCCTAGTGATCCTGGAAGCCAGTGTGAAA
GACATAGCTGCATTGAATTCCACGCTGGGTTGTTACCGCTACAGCTACGTGGGCCAGGGCCAGGTCCTCCGGCTG
AAGGGGCCTGACCACCTGGCCTCCAGCTGCCTGTGGCACCTGCAGGGCCCCAAGGACCTCATGCTCAAACTCCGG
CTGGAGTGGACGCTGGCAGAGTGCCGGGACCGACTGGCCATGTATGACGTGGCCGGGCCCCTGGAGAAGAGGCTC
ATCACCTCGGTGTACGGCTGCAGCCGCCAGGAGCCCGTGGTGGAGGTTCTGGCGTCGGGGGCCATCATGGCGGTC
GTCTGGAAGAAGGGCCTGCACAGCTACTACGACCCCTTCGTGCTCTCCGTGCAGCCGGTGGTCTTCCAGGCCTGT
GAAGTGAACCTGACGCTGGACAACAGGCTCGACTCCCAGGGCGTCCTCAGCACCCCGTACTTCCCCAGCTACTAC
TCGCCCCAAACCCACTGCTCCTGGCACCTCACGGTGCCCTCTCTGGACTACGGCTTGGCCCTCTGGTTTGATGCC
TATGCACTGAGGAGGCAGAAGTATGATTTGCCGTGCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGT
GGCTTGCGCATCCTGCAGCCCTACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACC
TCCCAGATCTCCCTCACCGGGCCCGGTGTGCGGGTGCACTATGGCTTGTACAACCAGTCGGACCCCTGCCCTGGA
GAGTTCCTCTGTTCTGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGGTCAAGGACTGCCCCAACGGCCTGGAT
GAGAGAAACTGCGTTTGCAGAGCCACATTCCAGTGCAAAGAGGACAGCACATGCATCTCACTGCCCAAGGTCTGT
GATGGGCAGCCTGATTGTCTCAACGGCAGCGATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTCACC
TTCCAGTGTGAGGACCGGAGCTGCGTGAAGAAGCCCAACCCGCAGTGTGATGGGCGGCCCGACTGCAGGGACGGC
TCGGATGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCCTCCAGCCGCATTGTTGGTGGAGCTGTGTCCTCCGAG
GGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTCGGGGTCGACACATCTGTGGGGGGCCCTCATCGCTGACCGC
TGGGTGATAACAGCTGCCTCCACTGCTTCCAGGAGGACAGCATGGCCTCCACGGTGCTGTGGACCGTGTTCCTGGGC
AAGGTGTGGCAGAACTCGCGCTGGCCTGGAGAGGTGTCCTTCAAGGTGAGCCGCCTGCTCCTGCACCCGTACCAC
GAAGAGGACAGCCATGACTACGACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTGCGCTCGGCCGCCGTGCGC
CCCGTCTGCCTGCCCGCGCGCTCCCACTTCTTCGAGCCCGGCCTGCACTGCTGGATTACGGGCTGGGGCGCCTTG
CGCGAGGGCGGCCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTGATCCCACAGGACCTGTGCAGCGAG
GCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGCCGGCTACCGCAAGGGCAAGAAGGATGCCTGTCAGGGT
GACTCAGGTGGTCCGCTGGTGTGCAAGGCACTCAGTGGCCGCTGGTTCCTGGCGGGGCTGGTCAGCTGGGGCCTG
GGCTGTGGCCGGCCTAACTACTTCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGTG
ACC**TGA**GGAACTGCCCCCCTGCAAAGCAGGGCCCACCTCCTGGACTCAGAGAGCCCAGGGCAACTGCCAAGCAGG
GGGACAAGTATTCTGGCGGGGGGTGGGGGAGAGAGCAGGCCCTGTGGTGGCAGGAGGTGGCATCTTGTCTCGTCC
CTGATGTCTGCTCCAGTGATGGCAGGAGGATGGAGAAGTGCCAGCAGCTGGGGGTCAAGACGTCCCCTGAGGACC
CAGGCCCACACCCAGCCCTTCTGCCTCCCAATTCTCTCTCCTCCGTCCCCTTCCTCCACTGCTGCCTAATGCAAG
GCAGTGGCTCAGCAGCAAGAATGCTGGTTCTACATCCCGAGGAGTGTCTGAGGTGCGCCCCACTCTGTACAGAGG
CTGTTTGGGCAGCCTTGCCTCCAGAGAGCAGATTCCAGCTTCGGAAGCCCCTGGTCTAACTTGGGATCTGGGAAT
GGAAGGTGCTCCCATCGGAGGGGACCCTCAGAGCCCTGGAGACTGCCAGGTGGGCCTGCTGCCACTGTAAGCCAA
AAGGTGGGGAAGTCCTGACTCCAGGGTCCTTGCCCCACCCCTGCCTGCCACCTGGGCCCTCACAGCCCAGACCCT
CACTGGGAGGTGAGCTCAGCTGCCCTTTGGAATAAAGCTGCCTGATCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 63

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49152

><subunit 1 of 1, 802 aa, 1 stop

><MW: 88846, pI: 6.41, NX(S/T): 7

MPVAEAPQVAGGQGDGGDGEEAEPEGMFKACEDSKRKARGYLRLVPLFVLLALLVLASAGVL

LWYFLGYKAEVMVSQVYSGSLRVLNRHFSQDLTRRESSAFRSETAKAQKMLKELITSTRLGT

YYNSSSVYSFGEGPLTCFFWFILQIPEHRRLMLSPEVVQALLVEELLSTVNSSAAVPYRAEY

EVDPEGLVILEASVKDIAALNSTLGCYRYSYVGQGQVLRLKGPDHLASSCLWHLQGPKDLML

KLRLEWTLAECRDRLAMYDVAGPLEKRLITSVYGCSRQEPVVEVLASGAIMAVVWKKGLHSY

YDPFVLSVQPVVFQACEVNLTLDNRLDSQGVLSTPYFPSYYSPQTHCSWHLTVPSLDYGLAL

WFDAYALRRQKYDLPCTQGQWTIQNRRLCGLRILQPYAERIPVVATAGITINFTSQISLTGP

GVRVHYGLYNQSDPCPGEFLCSVNGLCVPACDGVKDCPNGLDERNCVCRATFQCKEDSTCIS

LPKVCDGQPDCLNGSDEEQCQEGVPCGTFTFQCEDRSCVKKPNPQCDGRPDCRDGSDEEHCD

CGLQGPSSRIVGGAVSSEGEWPWQASLQVRGRHICGGALIADRWVITAAHCFQEDSMASTVL

WTVFLGKVWQNSRWPGEVSFKVSRLLLHPYHEEDSHDYDVALLQLDHPVVRSAAVRPVCLPA

RSHFFEPGLHCWITGWGALREGGPISNALQKVDVQLIPQDLCSEAYRYQVTPRMLCAGYRKG

KKDACQGDSGGPLVCKALSGRWFLAGLVSWGLGCGRPNYFGVYTRITGVISWIQQVVT

Important features:

Type II transmembrane domain:

amino acids 46-67

Serine proteases, trypsin family, histidine active site.

amino acids 604-609

N-glycosylation sites.

amino acids 127-130, 175-178, 207-210, 329-332, 424-427, 444-447
and 509-512

Kringle domains.

amino acids 746-758 and 592-609

Homologous region to Kallikrein Light Chain:

amino acids 568-779

Homologous region to Low-density lipoprotein receptor:

amino acids 451-567

FIGURE 64

GCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGTGGCTTGCGCATCCTGCAGCCC
TACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACCTCCCAGAT
CTCCCTCACCGGGCCCGGTGTGCGGGTGCACTATGGCTTGTACAACCAGTCGGACCCCTGCC
CTGGAGAGTTCCTCTGTTCTGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGGTCAAGGAC
TGCCCCAACGGCCTGGATGAGAGAAACTGCGTTTGCAGAGCCACATTCCAGTGCAAAGAGGA
CAGCACATGCATCTCACTGCCCAAGGTCTGTGATGGGCAGCCTGATTGTCTCAACGGCAGCG
ATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTCACCTTCCAGTGTGAGGACCGG
AGCTGCGTGAAGAAGCCCAACCCGCAGTGTGATGGGCGGCCCGACTGCAGGGACGGCTCGGA
TGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCCTCCAGCCGCATTGTTGGTGGAGCTGTGT
CCTCCGAGGGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTCGGGGTCGACACATCTGTGGG
GGGGCCCTCATCGCTGACCGCTGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCAT
GGCCTCCACGGTGCTGTGGACCGTGTTCCTGGGCAAGGTGTGGCAGAACTCGCGCTGGCCTG
GAGAGGTGTCCTTCAAGGTGAGCCGCCTGCTCCTGCACCCGTACCACGAAGAGGACAGCCAT
GACTACGACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTGCGCTCGGCCGCCGTGCGCCC
CGTCTGCCTGCCCGCGCGCTCCCACTTCTTCGAGCCCGGCCTGCACTGCTGGATTACGGGCT
GGGGCGCCTTGCGCGAGGGCGGCCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTG
ATCCCACAGGACCTGTGCAGCGAGGCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGC
CGGCTACCGCAAGGGCAAGAAGGATGCCTGTCAGGGTGACTCAGGTGGTCCGCTGGTGTGCA
AGGCACTCAGTGGCCGCTGGTTCCTGGCGGGGCTGGTCAGCTGGGGCCTGGGCTGTGGCCGG
CCTAACTACTTCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGT
GACCTGAGGAACTGCCCCCCTGCAAAGCAGGGCCCACCTCCTGGACTCAGAGAGCCCAGGGC
AACTGCCAAGCAGGGGGACAAGTAT

FIGURE 65

GGACGAGGGCAGATCTCGTTCTGGGGCAAGCCGTTGACACTCGCTCCCTGCCACCGCCCGGG

CTCCGTGCCGCCAAGTTTTCATTTTCCACCTTCTCTGCCTCCAGTCCCCCAGCCCCTGGCCG

AGAGAAGGGTCTTACCGGCCGGGATTGCTGGAAACACCAAGAGGTGGTTTTTGTTTTTTAAA

ACTTCTGTTTCTTGGGAGGGGGTGTGGCGGGGCAGG**ATG**AGCAACTCCGTTCCTCTGCTCTG

TTTCTGGAGCCTCTGCTATTGCTTTGCTGCGGGGAGCCCCGTACCTTTTGGTCCAGAGGGAC

GGCTGGAAGATAAGCTCCACAAACCCAAAGCTACACAGACTGAGGTCAAACCATCTGTGAGG

TTTAACCTCCGCACCTCCAAGGACCCAGAGCATGAAGGATGCTACCTCTCCGTCGGCCACAG

CCAGCCCTTAGAAGACTGCAGTTTCAACATGACAGCTAAAACCTTTTTCATCATTCACGGAT

GGACGATGAGCGGTATCTTTGAAAACTGGCTGCACAAACTCGTGTCAGCCCTGCACACAAGA

GAGAAGACGCCAATGTAGTTGTGGTTGACTGGCTCCCCCTGGCCCACCAGCTTTACACGGA

TGCGGTCAATAATACCAGGGTGGTGGGACACAGCATTGCCAGGATGCTCGACTGGCTGCAGG

AGAAGGACGATTTTTCTCTCGGGAATGTCCACTTGATCGGCTACAGCCTCGGAGCGCACGTG

GCCGGGTATGCAGGCAACTTCGTGAAAGGAACGGTGGGCCGAATCACAGGTTTGGATCCTGC

CGGGCCCATGTTTGAAGGGGCCGACATCCACAAGAGGCTCTCTCCGGACGATGCAGATTTTG

TGGATGTCCTCCACACCTACACGCGTTCCTTCGGCTTGAGCATTGGTATTCAGATGCCTGTG

GGCCACATTGACATCTACCCCAATGGGGGTGACTTCCAGCCAGGCTGTGGACTCAACGATGT

CTTGGGATCAATTGCATATGGAACAATCACAGAGGTGGTAAAATGTGAGCATGAGCGAGCCG

TCCACCTCTTTGTTGACTCTCTGGTGAATCAGGACAAGCCGAGTTTTGCCTTCCAGTGCACT

GACTCCAATCGCTTCAAAAAGGGGATCTGTCTGAGCTGCCGCAAGAACCGTTGTAATAGCAT

TGGCTACAATGCCAAGAAAATGAGGAACAAGAGGAACAGCAAAATGTACCTAAAAACCCGGG

CAGGCATGCCTTTCAGAGGTAACCTTCAGTCCCTGGAGTGTCCC**TGA**GGAAGGCCCTTAATA

CCTCCTTCTTAATACCATGCTGCAGAGCAGGGCACATCCTAGCCCAGGAGAAGTGGCCAGCA

CAATCCAATCAAATCGTTGCAAATCAGATTACACTGTGCATGTCCTAGGAAAGGGAATCTTT

ACAAAATAAACAGTGTGGACCCCTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAA

FIGURE 66

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49646

><subunit 1 of 1, 354 aa, 1 stop

><MW: 39362, pI: 8.35, NX(S/T): 2

MSNSVPLLCFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRTSKDPEHE

GCYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVVDWL

PLAHQLYTDAVNNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTV

GRITGLDPAGPMFEGADIHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDF

QPGCGLNDVLGSIAYGTITEVVKCEHERAVHLFVDSLVNQDKPSFAFQCTDSNRFKKGICLS

CRKNRCNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

Important features:

Signal peptide:

amino acids 1-16

Lipases, serine active site.

amino acids 163-172

N-glycosylation sites.

amino acids 80-83 and 136-139

FIGURE 67

```
CGGACGCGTGGGCGGACGCGTGGGCCTGGGCAAGGGCCGGGGCGCCGGGCCGAGCCACCTCTTCCCCTCCCCGC
TTCCCTGTCGCGCTCCGCTGGCTGGACGCGCTGGAGGAGTGGAGCAGCACCCGGCCGGCCCTGGGGGCTGACAGT
CGGCAAAGTTTGGCCCGAAGAGGAAGTGGTCTCAAACCCCGGCAGGTGGCGACCAGGCCAGACCAGGGGCGCTCG
CTGCCTGCGGGCGGGCTGTAGGCGAGGGCGCGCCCCAGTGCCGAGACCCGGGGCTTCAGGAGCCGGCCCCGGGAG
AGAAGAGTGCGGCGGCGGACGGAGAAAACAACTCCAAAGTTGGCGAAAGGCACCGCCCCTACTCCCGGGCTGCCG
CCGCCTCCCCGCCCCCAGCCCTGGCATCCAGAGTACGGGTCGAGCCCGGGCCATGGAGCCCCCCTGGGGAGGCGG
CACCAGGGAGCCTGGGCGCCCGGGGCTCCGCCGCGACCCCATCGGGTAGACCACAGAAGCTCCGGGACCCTTCCG
GCACCTCTGGACAGCCCAGG**ATG**CTGTTGGCCACCCTCCTCCTCCTCCTTGGAGGCGCTCTGGCCCATCCAG
ACCGGATTATTTTTCCAAATCATGCTTGTGAGGACCCCCCAGCAGTGCTCTTAGAAGTGCAGGGCACCTTACAGA
GGCCCCTGGTCCGGGACAGCCGCACCTCCCCTGCCAACTGCACCTGGCTCATCCTGGGCAGCAAGGAACAGACTG
TCACCATCAGGTTCCAGAAGCTACACCTGGCCTGTGGCTCAGAGCGCTTAACCCTACGCTCCCCTCTCCAGCCAC
TGATCTCCCTGTGTGAGGCACCTCCCAGCCCTCTGCAGCTGCCCGGGGGCAACGTCACCATCACTTACAGCTATG
CTGGGGCCAGAGCACCCATGGGCCAGGGCTTCCTGCTCTCCTACAGCCAAGATTGGCTGATGTGCCTGCAGGAAG
AGTTTCAGTGCCTGAACCACCGCTGTGTATCTGCTGTCCAGCGCTGTGATGGGGTTGATGCCTGTGGCGATGGCT
CTGATGAAGCAGGTTGCAGCTCAGACCCCTTCCCTGGCCTGACCCCAAGACCCGTCCCCTCCCTGCCTTGCAATG
TCACCTTGGAGGACTTCTATGGGGTCTTCTCCTCCTGGATATACACACCTAGCCTCAGTCTCCCACCCCCAGT
CCTGCCATTGGCTGCTGGACCCCCATGATGGCCGGCGGCTGGCCGTGCGCTTCACAGCCCTGGACTTGGGCTTTG
GAGATGCAGTGCATGTGTATGACGGCCCTGGGCCCCCTGAGAGCTCCCGACTACTGCGTAGTCTCACCCACTTCA
GCAATGGCAAGGCTGTCACTGTGGAGACACTGTCTGGCCAGGCTGTTGTGTCCTACCACACAGTTGCTTGGAGCA
ATGGTCGTGGCTTCAATGCCACCTACCATGTGCGGGGCTATTGCTTGCCTTGGGACAGACCCTGTGGCTTAGGCT
CTGGCCTGGGAGCTGGCGAAGGCCTAGGTGAGCGCTGCTACAGTGAGGCACAGCGCTGTGACGGCTCATGGGACT
GTGCTGACGGCACAGATGAGGAGGACTGCCCAGGCTGCCCACCTGGACACTTCCCCTGTGGGGCTGCTGGCACCT
CTGGTGCCACAGCCTGCTACCTGCCTGCTGACCGCTGCAACTACCAGACTTTCTGTGCTGATGGAGCAGATGAGA
GACGCTGTCGGCATTGCCAGCCTGGCAATTTCCGATGCCGGGACGAGAAGTGCGTGTATGAGACGTGGGTGTGCG
ATGGGCAGCCAGACTGTGCGGACGGCAGTGATGAGTGGGACTGCTCCTATGTTCTGCCCCGCCAAGGTCATTACAG
CTGCAGTCATTGGCAGCCTAGTGTGCGGCCTGCTCCTGGTCATCGCCCTGGGCTGCACCTGCAAGCTCTATGCCA
TTCGCACCCAGGAGTACAGCATCTTTGCCCCCCTCTCCCGGATGGAGGCTGAGATTGTGCAGCAGCAGGCACCCC
CTTCCTACGGGCAGCTCATTGCCCAGGGTGCCATCCCACCTGTAGAAGACTTTCCTACAGAGAATCCTAATGATA
ACTCAGTGCTGGGCAACCTGCGTTCTCTGCTACAGATCTTACGCCAGGATATGACTCCAGGAGGTGGCCCAGGTG
CCCGCCGTCGTCAGCGGGGCCGCTTGATGCGACGCCTGGTACGCCGTCTCCGCCGCTGGGGCTTGCTCCCTCGAA
CCAACACCCCGGCTCGGGCCTCTGAGGCCAGATCCCAGGTCACACCTTCTGCTGCTCCCCTTGAGGCCCTAGATG
GTGGCACAGGTCCAGCCCGTGAGGGCGGGGCAGTGGGTGGGCAAGATGGGGAGCAGGCACCCCCACTGCCCATCA
AGGCTCCCCTCCCATCTGCTAGCACGTCTCCAGCCCCCACTACTGTCCCTGAAGCCCCAGGGCCACTGCCCTCAC
TGCCCCTAGAGCCATCACTATTGTCTGGAGTGGTGCAGGCCCTGCGAGGCCGCCTGTTGCCCAGCCTGGGGCCCC
CAGGACCAACCCGGAGCCCCCCTGGACCCCACACAGCAGTCCTGGCCCTGGAAGATGAGGACGATGTGCTACTGG
TGCCACTGGCTGAGCCGGGGGTGTGGGTAGCTGAGGCAGAGGATGAGCCACTGCTTACC**TGA**GGGGACCTGGGGG
CTCTACTGAGGCCTCTCCCCTGGGGGCTCTACTCATAGTGGCACAACCTTTTAGAGGTGGGTCAGCCTCCCCTCC
ACCACTTCCTTCCCTGTCCCTGGATTTCAGGGACTTGGTGGGCCTCCCGTTGACCCTATGTAGCTGCTATAAAGT
TAAGTGTCCCTCAGGCAGGGAGAGGGCTCACAGAGTCTCCTCTGTACGTGGCCATGGCCAGACACCCCAGTCCCT
TCACCACCACCTGCTCCCCACGCCACCACCATTTGGGTGGCTGTTTTTAAAAAGTAAAGTTCTTAGAGGATCATA
GGTCTGGACACTCCATCCTTGCCAAACCTCTACCCAAAAGTGGCCTTAAGCACCGGAATGCCAATTAACTAGAGA
CCCTCCAGCCCCCAAGGGGAGGATTTGGGCAGAACCTGAGGTTTTGCCATCCACAATCCCTCCTACAGGGCCTGG
CTCACAAAAAGAGTGCAACAAATGCTTCTATTCCATAGCTACGGCATTGCTCAGTAAGTTGAGGTCAAAAATAAA
GGAATCATACATCTC
```

FIGURE 68

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631

<subunit 1 of 1, 713 aa, 1 stop

<MW: 76193, pI: 5.42, NX(S/T): 4

MLLATLLLLLLGGALAHPDRIIFPNHACEDPPAVLLEVQGTLQRPLVRDSRTSPANCTWLIL

GSKEQTVTIRFQKLHLACGSERLTLRSPLQPLISLCEAPPSPLQLPGGNVTITYSYAGARAP

MGQGFLLSYSQDWLMCLQEEFQCLNHRCVSAVQRCDGVDACGDGSDEAGCSSDPFPGLTPRP

VPSLPCNVTLEDFYGVFSSPGYTHLASVSHPQSCHWLLDPHDGRRLAVRFTALDLGFGDAVH

VYDGPGPPESSRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGFNATYHVRGYCLP

WDRPCGLGSGLGAGEGLGERCYSEAQRCDGSWDCADGTDEEDCPGCPPGHFPCGAAGTSGAT

ACYLPADRCNYQTFCADGADERRCRHCQPGNFRCRDEKCVYETWVCDGQPDCADGSDEWDCS

YVLPRKVITAAVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQQAPPSY

GQLIAQGAIPPVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGGGPGARRRQRGRLMRRLVR

RLRRWGLLPRTNTPARASEARSQVTPSAAPLEALDGGTGPAREGGAVGGQDGEQAPPLPIKA

PLPSASTSPAPTTVPEAPGPLPSLPLEPSLLSGVVQALRGRLLPSLGPPGPTRSPPGPHTAV

LALEDEDDVLLVPLAEPGVWVAEAEDEPLLT

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

FIGURE 69

CGAGCTGGGCGAGAAGTAGGGGAGGGCGGTGCTCCGCCGCGGTGGCGGTTGCTATCGCTTCG

CAGAACCTACTCAGGCAGCCAGCTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCA

GACGCG**ATG**GATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGTGAAAGG

CCACGTGAAGATGCTGCGGCTGGCACTAACTGTGACATCTATGACCTTTTTTATCATCGCAC

AAGCCCCTGAACCATATATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTTCATA

CTTTTATATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTTGGCCTTTGCTTGATAT

TATCAACTCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAG

AAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATGCTGTCTT

GCCGACGGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCCAGCGGTCCTTACCAGAAAAA

GCCTGTGCATGAAAAAAAGAAGTTTTG**TAA**TTTTATATTACTTTTTAGTTTGATACTAAGT

ATTAAACATATTTCTGTATTCTTCCAAAAAAAAAAAAAAAAAA

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645

><subunit 1 of 1, 152 aa, 1 stop

><MW: 17170, pI: 9.62, NX(S/T): 1

MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEPYIVITGFEVTVILFFILL
YVLRLDRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTLTVGGGVFALVTAVCCLAD
GALIYRKLLFNPSGPYQKKPVHEKKEVL

Important features:

Potential type II transmembrane domain:

amino acids 26-42

Other potential transmembrane domain:

amino acids 44-65, 81-101 and 109-129

Leucine zipper pattern

amino acids 78-99 and 85-106

N-myristoylation site.

amino acids 110-115

Ribonucleotide reductase large subunit protein

amino acids 116-127

FIGURE 71

GGGCGAGAAGTAGGGGAGGGCGTGTTCCGCCGCGGTGGCGGTTGCTATCGTTTTGCAGAACC

TACTCAGGCAGCCAGNTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCAGACGCGA

TGGATAACGTGCAGCCGAAAATAAAACATCGCCCCTTCTGCTTCAGTGTGAAAGGCCACGTG

AAGATGCTGCGGCTGGCACTAACTGNGACATCTATGACCTTTTTTATNATCGCACAAGCCCC

TGAACCATATATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTTCATACTTTTAT

ATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTTGGCCTTTGCTTGATATTATCAAC

TCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAGAAACCAC

AACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATGCTGTNTTGCCGAC

FIGURE 72

CAGCCCCGCGCGCCGGCCGAGTCGCTGAGCCGCGGCTGCCGGACGGGACGGGACCGGCTAGG

CTGGGCGCGCCCCCGGGCCCCGCCGTGGG**ATG**GGCGCACTGGCCCGGGCGCTGCTGCTGC

CTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTG

CCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCC

TGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCGCGG

GCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTAC

CTGGAGATGCTGATCGGGACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAG

TAACTTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGT

CTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACG

GGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAA

CATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATAC

TTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC

CTGGTGACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCC

CGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGT

ATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTG

AAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGC

CATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGG

AAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAG

CTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCT

GAGAGACGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGC

CCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCATCCACAAAT

GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAGAA

GAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTT

CCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTGAGCGAG

CCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTT

AATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCCCGTGACCCTGAGGTCGTCA

ATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA**TGA**ATAGCCAGGCCTGACCTCAAGCAA

CCATGAACTCAGCTATTAAGAAAATCACATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCG

CTTTCTCCTGTGCCCACCCGTCTTCAATCTCTGTTCTGCTCCCAGATGCCTTCTAGATTCAC

TGTCTTTTGATTCTTGATTTTCAAGCTTTCAAATCCTCCCTACTTCCAAGAAAAATAATTAA

AAAAAAAACTTCATTCTAA

FIGURE 73

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493

><subunit 1 of 1, 518 aa, 1 stop

><MW: 56180, pI: 5.08, NX(S/T): 2

MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLAL

ALEPALASPAGAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHS

YIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATIFESENF

FLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGS

LVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLR

LPQKVFDAVVEAVARASLIPEFSDGFWTGSQLACWTNSETPWSYFPKISIYLRDENSSRSFR

ITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA

EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLIVLLLLPFRC

QRRPRDPEVVNDESSLVRHRWK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 466-494

N-glycosylation sites.

amino acids 170-173 and 366-369

Leucine zipper pattern.

amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases

amino acids 109-118, 252-261 and 298-310

FIGURE 74

CGCCTCCGCCTTCGGAGGCTGACGCGCCCGGGCGCCGTTCCAGGCCTGTGCAGGGCGGATCG

GCAGCCGCCTGGCGGCGATCCAGGGCGGTGCGGGGCCTGGGCGGGAGCCGGGAGGCGCGGCC

GGC**ATG**GAGGCGCTGCTGCTGGGCGCGGGGTTGCTGCTGGGCGCTTACGTGCTTGTCTACTA

CAACCTGGTGAAGGCCCCGCCGTGCGGCGGCATGGGCAACCTGCGGGGCCGCACGGCCGTGG

TCACGGGCGCCAACAGCGGCATCGGAAAGATGACGGCGCTGGAGCTGGCGCGCCGGGGAGCG

CGCGTGGTGCTGGCCTGCCGCAGCCAGGAGCGCGGGGAGGCGGCTGCCTTCGACCTCCGCCA

GGAGAGTGGGAACAATGAGGTCATCTTCATGGCCTTGGACTTGGCCAGTCTGGCCTCGGTGC

GGGCCTTTGCCACTGCCTTTCTGAGCTCTGAGCCACGGTTGGACATCCTCATCCACAATGCC

GGTATCAGTTCCTGTGGCCGGACCCGTGAGGCGTTTAACCTGCTGCTTCGGGTGAACCATAT

CGGTCCCTTTCTGCTGACACATCTGCTGCTGCCTTGCCTGAAGGCATGTGCCCCTAGCCGCG

TGGTGGTGGTAGCCTCAGCTGCCCACTGTCGGGGACGTCTTGACTTCAAACGCCTGGACCGC

CCAGTGGTGGGCTGGCGGCAGGAGCTGCGGGCATATGCTGACACTAAGCTGGCTAATGTACT

GTTTGCCCGGGAGCTCGCCAACCAGCTTGAGGCCACTGGCGTCACCTGCTATGCAGCCCACC

CAGGGCCTGTGAACTCGGAGCTGTTCCTGCGCCATGTTCCTGGATGGCTGCGCCCACTTTTG

CGCCCATTGGCTTGGCTGGTGCTCCGGGCACCAAGAGGGGGTGCCCAGACACCCCTGTATTG

TGCTCTACAAGAGGGCATCGAGCCCCTCAGTGGGAGATATTTTGCCAACTGCCATGTGGAAG

AGGTGCCTCCAGCTGCCCGAGACGACCGGGCAGCCCATCGGCTATGGGAGGCCAGCAAGAGG

CTGGCAGGGCTTGGGCCTGGGGAGGATGCTGAACCCGATGAAGACCCCCAGTCTGAGGACTC

AGAGGCCCCATCTTCTCTAAGCACCCCCCACCCTGAGGAGCCCACAGTTTCTCAACCTTACC

CCAGCCCTCAGAGCTCACCAGATTTGTCTAAGATGACGCACCGAATTCAGGCTAAAGTTGAG

CCTGAGATCCAGCTCTCC**TAA**CCCTCAGGCCAGGATGCTTGCCATGGCACTTCATGGTCCTT

GAAAACCTCGGATGTGTGTGAGGCCATGCCCTGGACACTGACGGGTTTGTGATCTTGACCTC

CGTGGTTACTTTCTGGGGCCCCAAGCTGTGCCCTGGACATCTCTTTTCCTGGTTGAAGGAAT

AATGGGTGATTATTTCTTCCTGAGAGTGACAGTAACCCCAGATGGAGAGATAGGGGTATGCT

AGACACTGTGCTTCTCGGAAATTTGGATGTAGTATTTTCAGGCCCCACCCTTATTGATTCTG

ATCAGCTCTGGAGCAGAGGCAGGGAGTTTGCAATGTGATGCACTGCCAACATTGAGAATTAG

TGAACTGATCCCTTTGCAACCGTCTAGCTAGGTAGTTAAATTACCCCCATGTTAATGAAGCG

GAATTAGGCTCCCGAGCTAAGGGACTCGCCTAGGGTCTCACAGTGAGTAGGAGGAGGGCCTG

GGATCTGAACCCAAGGGTCTGAGGCCAGGGCCGACTGCCGTAAGATGGGTGCTGAGAAGTGA

GTCAGGGCAGGGCAGCTGGTATCGAGGTGCCCCATGGGAGTAAGGGGACGCCTTCCGGGCGG

ATGCAGGGCTGGGGTCATCTGTATCTGAAGCCCCTCGGAATAAAGCGCGTTGACCGCCAAAA

AAAAAAAAAAAAAAAAAA

FIGURE 75

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48227

<subunit 1 of 1, 377 aa, 1 stop

<MW: 40849, pI: 7.98, NX(S/T): 0

MEALLLGAGLLLGAYVLVYYNLVKAPPCGGMGNLRGRTAVVTGANSGIGKMTALELARRGAR

VVLACRSQERGEAAAFDLRQESGNNEVIFMALDLASLASVRAFATAFLSSEPRLDILIHNAG

ISSCGRTREAFNLLLRVNHIGPFLLTHLLLPCLKACAPSRVVVVASAAHCRGRLDFKRLDRP

VVGWRQELRAYADTKLANVLFARELANQLEATGVTCYAAHPGPVNSELFLRHVPGWLRPLLR

PLAWLVLRAPRGGAQTPLYCALQEGIEPLSGRYFANCHVEEVPPAARDDRAAHRLWEASKRL

AGLGPGEDAEPDEDPQSEDSEAPSSLSTPHPEEPTVSQPYPSPQSSPDLSKMTHRIQAKVEP

EIQLS

Important features:

Signal peptide:

amino acids 1-16

Glycosaminoglycan attachment site.

amino acids 46-49

Short-chain alcohol dehydrogenase family

amino acids 37-49 and 114-124

FIGURE 76

GGAGGAGACAGCCTCCTGGGGGGCAGGGGTTCCCTGCCTCTGCTGCTCCTGCTCATC**<u>ATG</u>**GGAGGCATGGCTCAG
GACTCCCCGCCCCAGATCCTAGTCCACCCCCAGGACCAGCTGTTCCAGGGCCCTGGCCCTGCCAGGATGAGCTGC
CAAGCCTCAGGCCAGCCACCTCCCACCATCCGCTGGTTGCTGAATGGGCAGCCCCTGAGCATGGTGCCCCCAGAC
CCACACCACCTCCTGCCTGATGGGACCCTTCTGCTGCTACAGCCCCCTGCCCGGGGACATGCCCACGATGGCCAG
GCCCTGTCCACAGACCTGGGTGTCTACACATGTGAGGCCAGCAACCGGCTTGGCACGGCAGTCAGCAGAGGCGCT
CGGCTGTCTGTGGCTGTCCTCCGGGAGGATTTCCAGATCCAGCCTCGGGACATGGTGGCTGTGGTGGGTGAGCAG
TTTACTCTGGAATGTGGGCCGCCCTGGGGCCACCCAGAGCCCACAGTCTCATGGTGGAAAGATGGGAAACCCCTG
GCCCTCCAGCCCGGAAGGCACACAGTGTCCGGGGGTCCCTGCTGATGGCAAGAGCAGAGAAGAGTGACGAAGGG
ACCTACATGTGTGTGTGGCCACCAACAGCGCAGGACATAGGGAGAGCCGCGCAGCCCGGGTTTCCATCCAGGAGCCC
CAGGACTACACGGAGCCTGTGGAGCTTCTGGCTGTGCGAATTCAGCTGGAAAATGTGACACTGCTGAACCCGGAT
CCTGCAGAGGGCCCCAAGCCTAGACCGGCGGTGTGGCTCAGCTGGAAGGTCAGTGGCCCTGCTGCGCCTGCCCAA
TCTTACACGGCCTTGTTCAGGACCCAGACTGCCCCGGGAGGCCAGGGAGCTCCGTGGGCAGAGGAGCTGCTGGCC
GGCTGGCAGAGCGCAGAGCTTGGAGGCCTCCACTGGGGCCAAGACTACGAGTTCAAAGTGAGACCATCCTCTGGC
CGGGCTCGAGGCCCTGACAGCAACGTGCTGCTCCTGAGGCTGCCGGAAAAAGTGCCCAGTGCCCCACCTCAGGAA
GTGACTCTAAAGCCTGGCAATGGCACTGTCTTTGTGAGCTGGGTCCCACCACCTGCTGAAAACCACAATGGCATC
ATCCGTGGCTACCAGGTCTGGAGCCTGGGCAACACATCACTGCCACCAGCCAACTGGACTGTAGTTGGTGAGCAG
ACCCAGCTGGAAATCGCCACCCATATGCCAGGCTCCTACTGCGTGCAAGTGGCTGCAGTCACTGGTGCTGGAGCT
GGGGAGCCCAGTAGACCTGTCTGCCTCCTTTTAGAGCAGGCCATGGAGCGAGCCACCCAAGAACCCAGTGAGCAT
GGTCCCTGGACCCTGGAGCAGCTGAGGGCTACCTTGAAGCGGCCTGAGGTCATTGCCACCTGCGGTGTTGCACTC
TGGCTGCTGCTTCTGGGCACCGCCGTGTGTATCCACCGCCGGCGCCGAGCTAGGGTGCACCTGGGCCCAGGTCTG
TACAGATATACCAGTGAGGATGCCATCCTAAAACACAGGATGGATCACAGTGACTCCCAGTGGTTGGCAGACACT
TGGCGTTCCACCTCTGGCTCTCGGGACCTGAGCAGCAGCAGCAGCCTCAGCAGTCGGCTGGGGGCGGATGCCCGG
GACCCACTAGACTGTCGTCGCTCCTTGCTCTCCTGGGACTCCCGAAGCCCCGGCGTGCCCCTGCTTCCAGACACC
AGCACTTTTTATGGCTCCCTCATCGCTGAGCTGCCCTCCAGTACCCCAGCCAGGCCAAGTCCCCAGGTCCCAGCT
GTCAGGCGCCTCCCACCCCAGCTGGCCCAGCTCTCCAGCCCCTGTTCCAGCTCAGACAGCCTCTGCAGCCGCAGG
GGACTCTCTTCTCCCCGCTTGTCTCTGGCCCCTGCAGAGGCTTGGAAGGCCAAAAAGAAGCAGGAGCTGCAGCAT
GCCAACAGTTCCCCACTGCTCCGGGGCAGCCACTCCTTGGAGCTCCGGGCCTGTGAGTTAGGAAATAGAGGTTCC
AAGAACCTTTCCCAAAGCCCAGGAGCTGTGCCCCAAGCTCTGGTTGCCTGGCGGGCCCTGGGACCGAAACTCCTC
AGCTCCTCAAATGAGCTGGTTACTCGTCATCTCCCTCCAGCACCCCTCTTTCCTCATGAAACTCCCCCAACTCAG
AGTCAACAGACCCAGCCTCCGGTGGCACCACAGGCTCCCTCCTCCATCCTGCTGCCAGCAGCCCCCATCCCCATC
CTTAGCCCCTGCAGTCCCCCTAGCCCCCAGGCCTCTTCCCTCTCTGGCCCCAGCCCAGCTTCCAGTCGCCTGTCC
AGCTCCTCACTGTCATCCCTGGGGGAGGATCAAGACAGCGTGCTGACCCCTGAGGAGGTAGCCCTGTGCTTGGAA
CTCAGTGAGGGTGAGGAGACTCCCAGGAACAGCGTCTCTCCCATGCCAAGGGCTCCTTCACCCCCCACCACCTAT
GGGTACATCAGCGTCCCAACAGCCTCAGAGTTCACGGACATGGGCAGGACTGGAGGAGGGGTGGGGCCCAAGGGG
GGAGTCTTGCTGTGCCCACCTCGGCCCTGCCTCACCCCCACCCCCAGCGAGGGCTCCTTAGCCAATGGTTGGGGC
TCAGCCTCTGAGGACAATGCCGCCAGCGCCAGAGCCAGCCTTGTCAGCTCCTCCGATGGCTCCTTCCTCGCTGAT
GCTCACTTTGCCCGGGCCCTGGCAGTGGCTGTGGATAGCTTTGGTTTCGGTCTAGAGCCCAGGGAGGCAGACTGC
GTCTTCATAGATGCCTCATCACCTCCCTCCCCACGGGATGAGATCTTCCTGACCCCCAACCTCTCCCTGCCCCTG
TGGGAGTGGAGGCCAGACTGGTTGGAAGACATGGAGGTCAGCCACACCCAGCGGCTGGGAAGGGGGATGCCTCCC
TGGCCCCCTGACTCTCAGATCTCTTCCCAGAGAAGTCAGCTCCACTGTCGTATGCCCAAGGCTGGTGCTTCTCCT
GTAGATTACTCC**<u>TGA</u>**ACCGTGTCCCTGAGACTTCCCAGACGGGAATCAGAACCACTTCTCCTGTCCACCCACAAG
ACCTGGGCTGTGGTGTGTGGGTCTTGGCCTGTGTTTCTCTGCAGCTGGGGTCCACCTTCCCAAGCCTCCAGAGAG
TTCTCCCTCCACGATTGTGAAAACAAATGAAAACAAAATTAGAGCAAAGCTGACCTGGAGCCCTCAGGGAGCAAA
ACATCATCTCCACCTGACTCCTAGCCACTGCTTTCTCCTCTGTGCCATCCACTCCCACCACCAGGTTGTTTTGGC
CTGAGGAGCAGCCCTGCCTGCTGCTCTTCCCCCACCATTTGGATCACAGGAAGTGGAGGAGCCAGAGGTGCCTTT
GTGGAGGACAGCAGTGGCTGCTGGGAGAGGGCTGTGGAGGAAGGAGCTTCTCGGAGCCCCCTCTCAGCCTTACCT
GGGCCCCTCCTCTAGAGAAGAGCTCAACTCTCTCCCAACCTCACCATGGAAAGAAAATAATTATGAATGCCACTG
AGGCATGAGGCCCTACCTCATGCCAAACAAAGGGTTCAAGGCTGGGTCTAGCGAGGATGCTGAAGGAAGGGAGG
TATGAGACCGTAGGTCAAAAGCACCATCCTCGTACTGTTGTCACTATGAGCTTAAGAAATTTGATACCATAAAAT
GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 77

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404

<subunit 1 of 1, 985 aa, 1 stop

<MW: 105336, pI: 6.55, NX(S/T): 7

MGGMAQDSPPQILVHPQDQLFQGPGPARMSCQASGQPPPTIRWLLNGQPLSMVPPDPHHLLP

DGTLLLLQPPARGHAHDGQALSTDLGVYTCEASNRLGTAVSRGARLSVAVLREDFQIQPRDM

VAVVGEQFTLECGPPWGHPEPTVSWWKDGKPLALQPGRHTVSGGSLLMARAEKSDEGTYMCV

ATNSAGHRESRAARVSIQEPQDYTEPVELLAVRIQLENVTLLNPDPAEGPKPRPAVWLSWKV

SGPAAPAQSYTALFRTQTAPGGQGAPWAEELLAGWQSAELGGLHWGQDYEFKVRPSSGRARG

PDSNVLLLRLPEKVPSAPPQEVTLKPGNGTVFVSWVPPPAENHNGIIRGYQVWSLGNTSLPP

ANWTVVGEQTQLEIATHMPGSYCVQVAAVTGAGAGEPSRPVCLLLEQAMERATQEPSEHGPW

TLEQLRATLKRPEVIATCGVALWLLLLGTAVCIHRRRRARVHLGPGLYRYTSEDAILKHRMD

HSDSQWLADTWRSTSGSRDLSSSSSLSSRLGADARDPLDCRRSLLSWDSRSPGVPLLPDTST

FYGSLIAELPSSTPARPSPQVPAVRRLPPQLAQLSSPCSSSDSLCSRRGLSSPRLSLAPAEA

WKAKKKQELQHANSSPLLRGSHSLELRACELGNRGSKNLSQSPGAVPQALVAWRALGPKLLS

SSNELVTRHLPPAPLFPHETPPTQSQQTQPPVAPQAPSSILLPAAPIPILSPCSPPSPQASS

LSGPSPASSRLSSSSLSSLGEDQDSVLTPEEVALCLELSEGEETPRNSVSPMPRAPSPPTTY

GYISVPTASEFTDMGRTGGGVGPKGGVLLCPPRPCLTPTPSEGSLANGWGSASEDNAASARA

SLVSSSDGSFLADAHFARALAVAVDSFGFGLEPREADCVFIDASSPPSPRDEIFLTPNLSLP

LWEWRPDWLEDMEVSHTQRLGRGMPPWPPDSQISSQRSQLHCRMPKAGASPVDYS

Important features:

Transmembrane domain:

amino acids 448-467

N-glycosylation sites:

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

N-myristoylation sites.

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

Phosphotyrosine interaction domain proteins

amino acids 740-753

FIGURE 78

CTCCCACGGTGTCCAGCGCCCAGA**ATG**CGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT
CCCAGGTTATGAAGCCCTGGAGGGCCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT
CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT
GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA
ACCTCACCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG
TCTTTACTGATCTCTCTGTTCGTCTTTCCAGGACCCTGCTGTCCTCCCTCCCCTTCTCCCAC
CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCCC
CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG
GCTGAGGCCCCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCATGCAGC
TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG
GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCCTGGTGCTGCTGAGCCTTCTGTCAGC
CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA
CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC
CCTTCCCAGGCCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGA
GCTGGGCTTCTCGAAGTTTGTCTCAGCG**TAG**GGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT
GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCCACCTCAGCCTCAGAG
TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCCAGGCTCTCCTCTTGCATGTTCCA
GCCTGACCTAGAAGCGTTTGTCAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG
GAGACTGGGACATCCCTGATAGGTTCACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA
GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACTCCTGGGC
CTCATGCCCAGTGTCGGACCCTGCCTTCCTCCCACTCCAGACCCCACCTTGTCTTCCCTCCC
TGGCGTCCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT
GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT
GTGAAAAACGTGATTCCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG
GACTCTGAATTCTAACAATGCCCAGTGACTGTCGCACTTGAGTTTGAGGGCCAGTGGGCCTG
ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC
CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG
TCCAGGCCTTGGTCAGGTCAGGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG
TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCCTTTGGAAAAAATGATGAAGA
AAACCTTGGCTCCTTCCTTGTCTGGAAAGGGTTACTTGCCTATGGGTTCTGGTGGCTAGAGA
GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG
ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTCGGGGGTGGTGGTAAAGTA
GCACAACTACTATTTTTTTTCTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT
GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACTCCGCCTCCTGGGTTCAAGTGATT
CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATT
TTTGTACTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGAC
CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG
TCTGGCCCTATTTCCTTTAAAAAGTGAAATTAAGAGTTGTTCAGTATGCAAAACTTGGAAAG
ATGGAGGAGAAAAGAAAAGGAAGAAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT
TATTTCGTTTTGTTGTACTTCCTTCCACTCTTTTCTTCTTCACATAATTTGCCGGTGTTCTT
TTTACAGAGCAATTATCTTGTATATACAACTTTGTATCCTGCCTTTTCCACCTTATCGTTCC
ATCACTTTATTCCAGCACTTCTCTGTGTTTTACAGACCTTTTTATAAATAAAATGTTCATCA
GCTGCATAAAAAAAAAAAAAA

FIGURE 79

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS

GTIYAEEEGQETMKGRVSIRDSRQELSLIVTLWNLTLQDAGEYWCGVEKRGPDESLLISLFV

FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG

TSQYGHERTSQYTGTSPHPATSPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI

LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAEEKEAPSQAPEGD

VISMPPLHTSEEELGFSKFVSA

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

FIGURE 80

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGA
GCCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTC
GCTCTGTTGCCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCT
GGACTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAA
GCCCTGTTTCTTCTCCTTCTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTCATCCCAA
AGCTCAGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAG
CAGGGCTCTCAGAAGGCGGTGGTGCCCAGCTGGGATC**ATG**TTGTTGGCCCTGGTCTGTCTGC
TCAGCTGCCTGCTACCCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGGCCAGAGTG
CTACATGACTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGC
TTATTTCACAAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACA
ACGGGATCTTCCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCCGAACGTCCCCAAC
GTGTGCCGGATGTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGC
CATGAAGATAACCCAAGAGCCTCAGGGTCTGGGTTACTGGGAGGCCTGGAGGCATCACTGCC
AGGGAAAAGACCTCACTGAATGGGTGGATGGCTGTGACTTC**TAG**GATGGACGGAACCATGCA
CAGCAGGCTGGGAAATGTGGTTTGGTTCCTGACCTAGGCTTGGGAAGACAAGCCAGCGAATA
AAGGATGGTTGAACGTGAAA

FIGURE 81

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187

<subunit 1 of 1, 146 aa, 1 stop

<MW: 16430, pI: 5.05, NX(S/T): 1

MLLALVCLLSCLLPSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALD
YEADGSTNNGIFQINSRRWCSNLTPNVPNVCRMYCSDLLNPNLKDTVICAMKITQEPQGLGY
WEAWRHHCQGKDLTEWVDGCDF

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homolgous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

FIGURE 82

AGCCGCTGCCCCGGGCCGGGCGCCCGCGGCGGCACC**ATG**AGTCCCCGCTCGTGCCTGCGTTC

GCTGCGCCTCCTCGTCTTCGCCGTCTTCTCAGCCGCCGCGAGCAACTGGCTGTACCTGGCCA

AGCTGTCGTCGGTGGGGAGCATCTCAGAGGAGGAGACGTGCGAGAAACTCAAGGGCCTGATC

CAGAGGCAGGTGCAGATGTGCAAGCGGAACCTGGAAGTCATGGACTCGGTGCGCCGCGGTGC

CCAGCTGGCCATTGAGGAGTGCCAGTACCAGTTCCGGAACCGGCGCTGGAACTGCTCCACAC

TCGACTCCTTGCCCGTCTTCGGCAAGGTGGTGACGCAAGGGACTCGGGAGGCGGCCTTCGTG

TACGCCATCTCTTCGGCAGGTGTGGCCTTTGCAGTGACGCGGGCGTGCAGCAGTGGGGAGCT

GGAGAAGTGCGGCTGTGACAGGACAGTGCATGGGGTCAGCCCACAGGGCTTCCAGTGGTCAG

GATGCTCTGACAACATCGCCTACGGTGTGGCCTTCTCACAGTCGTTTGTGGATGTGCGGGAG

AGAAGCAAGGGGGCCTCGTCCAGCAGAGCCCTCATGAACCTCCACAACAATGAGGCCGGCAG

GAAGGCCATCCTGACACACATGCGGGTGGAATGCAAGTGCCACGGGGTGTCAGGCTCCTGTG

AGGTAAAGACGTGCTGGCGAGCCGTGCCGCCCTTCCGCCAGGTGGGTCACGCACTGAAGGAG

AAGTTTGATGGTGCCACTGAGGTGGAGCCACGCCGCGTGGGCTCCTCCAGGGCACTGGTACC

ACGCAACGCACAGTTCAAGCCGCACACAGATGAGGACCTGGTGTACTTGGAGCCTAGCCCCG

ACTTCTGTGAGCAGGACATGCGCAGCGGCGTGCTGGGCACGAGGGGCCGCACATGCAACAAG

ACGTCCAAGGCCATCGACGGCTGTGAGCTGCTGTGCTGTGGCCGCGGCTTCCACACGGCGCA

GGTGGAGCTGGCTGAACGCTGCAGCTGCAAATTCCACTGGTGCTGCTTCGTCAAGTGCCGGC

AGTGCCAGCGGCTCGTGGAGTTGCACACGTGCCGA**TGA**CCGCCTGCCTAGCCCTGCGCCGGC

AACCACCTAGTGGCCCAGGGAAGGCCGATAATTTAAACAGTCTCCCACCACCTACCCCAAGA

GATACTGGTTGTATTTTTTGTTCTGGTTTGGTTTTTGGGTCCTCATGTTATTTATTGCCGAA

ACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGCCTCCCCAAAGCCTGGGCCTTTGTGGCT

GCCACTGACCAAAGGGACCTTGCTCGTGCCGCTGGCTGCCCGCATGTGGCTGCCACTGACCA

CTCAGTTGTTATCTGTGTCCGTTTTTCTACTTGCAGACCTAAGGTGGAGTAACAAGGAGTAT

TACCACCACATGGCTACTGACCGTGTCATCGGGGAAGAGGGGGCCTTATGGCAGGGAAAATA

GGTACCGACTTGATGGAAGTCACACCCTCTGGAAAAAAGAACTCTTAACTCTCCAGCACACA

TACACATGGACTCCTGGCAGCTTGAGCCTAGAAGCCATGTCTCTCAAATGCCCTGAGAAAGG

GAACAAGCAGATACCAGGTCAAGGGCACCAGGTTCATTTCAGCCCTTACATGGACAGCTAGA

GGTTCGATATCTGTGGGTCCTTCCAGGCAAGAAGAGGGAGATGAGAGCAAGAGACGACTGAA

GTCCCACCCTAGAACCCAGCCTGCCCCAGCCTGCCCCTGGGAAGAGGAAACTTAACCACTCC

CCAGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACCAGGGATCCCGGCTGTGCCTT

TGCAGTCATGCCCGAGTCACCTTTCACAGCGCTGTTCCTCCATGAAACTGAAAAACACACAC

ACACACACACACACACACACACACACACACACACGGACACACACACACCTGCGAGA

GAGAGGGAGGAAAGGGCTGTGCCTTTGCAGTCATGCCCGAGTCACCTTTCACAGCACTGTTCCTC

FIGURE 83

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48328

<subunit 1 of 1, 351 aa, 1 stop

<MW: 39052, pI: 8.97, NX(S/T): 2

MSPRSCLRSLRLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLKGLIQRQVQMCKRNLE

VMDSVRRGAQLAIEECQYQFRNRRWNCSTLDSLPVFGKVVTQGTREAAFVYAISSAGVAFAV

TRACSSGELEKCGCDRTVHGVSPQGFQWSGCSDNIAYGVAFSQSFVDVRERSKGASSSRALM

NLHNNEAGRKAILTHMRVECKCHGVSGSCEVKTCWRAVPPFRQVGHALKEKFDGATEVEPRR

VGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTSKAIDGCELLC

CGRGFHTAQVELAERCSCKFHWCCFVKCRQCQRLVELHTCR

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 88-91 and 297-300

Wnt-1 family signature.

amino acids 206-215

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

FIGURE 84

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGTGCCTGCAT

CGCC**ATG**GACACCACCAGGTACAGCAAGTGGGGCGGCAGCTCCGAGGAGGTCCCCGGAGGGC

CCTGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGGCCCTGGCTGTCCTGGTC

ACCACAGTCCTTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGC

GGCGCTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGGCGCTGG

GTGCCCTGAAGGAGGAGGTCGGAGACTGCCACAGCTGCTGCTCGGGGACGCAGGCGCAGCTG

CAGACCACGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCT

GCGGGAACTGCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGGCCGTGAGGACG

TCCGCACTGAGCTGTTCCGGGCGCTGGAGGCCGTGAGGCTCCAGAACAACTCCTGCGAGCCG

TGCCCCACGTCGTGGCTGTCCTTCGAGGGCTCCTGCTACTTTTTCTCTGTGCCAAAGACGAC

GTGGGCGGCGGCGCAGGATCACTGCGCAGATGCCAGCGCGCACCTGGTGATCGTTGGGGGCC

TGGATGAGCAGGGCTTCCTCACTCGGAACACGCGTGGCCGTGGTTACTGGCTGGGCCTGAGG

GCTGTGCGCCATCTGGGCAAGGTTCAGGGCTACCAGTGGGTGGACGGAGTCTCTCTCAGCTT

CAGCCACTGGAACCAGGGAGAGCCCAATGACGCTTGGGGGCGCGAGAACTGTGTCATGATGC

TGCACACGGGGCTGTGGAACGACGCACCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAG

AAAAGGCACAACTGC**TGA**CCCCGCCCAGTGCCCTGGAGCCGCGCCCATTGCAGCATGTCGTA

TCCTGGGGGCTGCTCACCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTTTCTTCCT

CATCCACCGCTGCTGAGTCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCCAGTGCC

TGGGCTCTGGGACCTCCATGCCGACCTCATCCTAACTCCACTCACGCAGACCCAACCTAACC

TCCACTAGCTCCAAAATCCCTGCTCCTGCGTCCCCGTGATATGCCTCCACTTCTCTCCCTAA

CCAAGGTTAGGTGACTGAGGACTGGAGCTGTTTGGTTTTCTCGCATTTTCCACCAAACTGGA

AGCTGTTTTTGCAGCCTGAGGAAGCATCAATAAATATTTGAGAAATGAAAAAA

FIGURE 85

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56352

<subunit 1 of 1, 293 aa, 1 stop

<MW: 32562, pI: 6.53, NX(S/T): 2

MDTTRYSKWGGSSEEVPGGPWGRWVHWSRRPLFLALAVLVTTVLWAVILSILLSKASTERAA

LLDGHDLLRTNASKQTAALGALKEEVGDCHSCCSGTQAQLQTTRAELGEAQAKLMEQESALR

ELRERVTQGLAEAGRGREDVRTELFRALEAVRLQNNSCEPCPTSWLSFEGSCYFFSVPKTTW

AAAQDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSFS

HWNQGEPNDAWGRENCVMMLHTGLWNDAPCDSEKDGWICEKRHNC

Important features:

Type II transmembrane domain:

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

FIGURE 86

GCCAGGGGAAGAGGGTGATCCGACCCGGGGAAGGTCGCTGGGCAGGGCGAGTTGGGAAAGCG

GCAGCCCCGCCGCCCCGCAGCCCCTTCTCCTCCTTTCTCCCACGTCCTATCTGCCTCTCG

CTGGAGGCCAGGCCGTGCAGCATCGAAGACAGGAGGAACTGGAGCCTCATTGGCCGGCCCGG

GGCGCCGGCCTCGGGCTTAAATAGGAGCTCCGGGCTCTGGCTGGGACCCGACCGCTGCCGGC

CGCGCTCCCGCTGCTCCTGCCGGGTG**ATG**GAAAACCCCAGCCCGGCCGCCGCCCTGGGCAAG

GCCCTCTGCGCTCTCCTCCTGGCCACTCTCGGCGCCGCCGGCCAGCCTCTTGGGGGAGAGTC

CATCTGTTCCGCCAGAGCCCCGGCCAAATACAGCATCACCTTCACGGGCAAGTGGAGCCAGA

CGGCCTTCCCCAAGCAGTACCCCCTGTTCCGCCCCCCTGCGCAGTGGTCTTCGCTGCTGGGG

GCCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACCAGTACGTCAGTAACGGGCTGCG

CGACTTTGCGGAGCGCGGCGAGGCCTGGGCGCTGATGAAGGAGATCGAGGCGGCGGGGGAGG

CGCTGCAGAGCGTGCACGAGGTGTTTTCGGCGCCCGCCGTCCCCAGCGGCACCGGGCAGACG

TCGGCGGAGCTGGAGGTGCAGCGCAGGCACTCGCTGGTCTCGTTTGTGGTGCGCATCGTGCC

CAGCCCCGACTGGTTCGTGGGCGTGGACAGCCTGGACCTGTGCGACGGGGACCGTTGGCGGG

AACAGGCGGCGCTGGACCTGTACCCCTACGACGCCGGGACGGACAGCGGCTTCACCTTCTCC

TCCCCCAACTTCGCCACCATCCCGCAGGACACGGTGACCGAGATAACGTCCTCCTCTCCCAG

CCACCCGGCCAACTCCTTCTACTACCCGCGGCTGAAGGCCCTGCCTCCCATCGCCAGGGTGA

CACTGCTGCGGCTGCGACAGAGCCCCAGGGCCTTCATCCCTCCCGCCCCAGTCCTGCCCAGC

AGGGACAATGAGATTGTAGACAGCGCCTCAGTTCCAGAAACGCCGCTGGACTGCGAGGTCTC

CCTGTGGTCGTCCTGGGGACTGTGCGGAGGCCACTGTGGGAGGCTCGGGACCAAGAGCAGGA

CTCGCTACGTCCGGGTCCAGCCCGCCAACAACGGGAGCCCCTGCCCCGAGCTCGAAGAAGAG

GCTGAGTGCGTCCCTGATAACTGCGTC**TAA**GACCAGAGCCCCGCAGCCCCTGGGGCCCCCCG

GAGCCATGGGGTGTCGGGGGCTCCTGTGCAGGCTCATGCTGCAGGCGGCCGAGGGCACAGGG

GGTTTCGCGCTGCTCCTGACCGCGGTGAGGCCGCGCCGACCATCTCTGCACTGAAGGGCCCT

CTGGTGGCCGGCACGGGCATTGGGAAACAGCCTCCTCCTTTCCCAACCTTGCTTCTTAGGGG

CCCCCGTGTCCCGTCTGCTCTCAGCCTCCTCCTCCTGCAGGATAAAGTCATCCCCAAGGCTC

CAGCTACTCTAAATTATGTCTCCTTATAAGTTATTGCTGCTCCAGGAGATTGTCCTTCATCG

TCCAGGGGCCTGGCTCCCACGTGGTTGCAGATACCTCAGACCTGGTGCTCTAGGCTGTGCTG

AGCCCACTCTCCCGAGGGCGCATCCAAGCGGGGGCCACTTGAGAAGTGAATAAATGGGGCGG

TTTCGGAAGCGTCAGTGTTTCCATGTTATGGATCTCTCTGCGTTTGAATAAAGACTATCTCT

GTTGCTCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 87

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53971

><subunit 1 of 1, 331 aa, 1 stop

><MW: 35844, pI: 5.45, NX(S/T): 2

MENPSPAAALGKALCALLLATLGAAGQPLGGESICSARAPAKYSITFTGKWSQTAFPKQYPL

FRPPAQWSSLLGAAHSSDYSMWRKNQYVSNGLRDFAERGEAWALMKEIEAAGEALQSVHEVF

SAPAVPSGTGQTSAELEVQRRHSLVSFVVRIVPSPDWFVGVDSLDLCDGDRWREQAALDLYP

YDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSHPANSFYYPRLKALPPIARVTLLRLRQSP

RAFIPPAPVLPSRDNEIVDSASVPETPLDCEVSLWSSWGLCGGHCGRLGTKSRTRYVRVQPA

NNGSPCPELEEEAECVPDNCV

Important features:

Signal peptide:

amino acids 1-26

FIGURE 88

```
GGCGGCGTCCGTGAGGGGCTCCTTTGGGCAGGGGTAGTGTTTGGTGTCCCTGTCTTGCGTGA
TATTGACAAACTGAAGCTTTCCTGCACCACTGGACTTAAGGAAGAGTGTACTCGTAGGCGGA
CAGCTTTAGTGGCCGGCCGGCCGCTCTCATCCCCCGTAAGGAGCAGAGTCCTTTGTACTGAC
CAAGATGAGCAACATCTACATCCAGGAGCCTCCCACGAATGGGAAGGTTTTATTGAAAACTA
CAGCTGGAGATATTGACATAGAGTTGTGGTCCAAAGAAGCTCCTAAAGCTTGCAGAAATTTT
ATCCAACTTTGTTTGGAAGCTTATTATGACAATACCATTTTTCATAGAGTTGTGCCTGGTTT
CATAGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGTGGAGAGTCTATCTATGGAGCGC
CATTCAAAGATGAATTTCATTCACGGTTGCGTTTTAATCGGAGAGGACTGGTTGCCATGGCA
AATGCTGGTTCTCATGATAATGGCAGCCAGTTTTTCTTCACACTGGGTCGAGCAGATGAACT
TAACAATAAGCATACCATCTTTGGAAAGGTTACAGGGGATACAGTATATAACATGTTGCGAC
TGTCAGAAGTAGACATTGATGATGACGAAAGACCACATAATCCACACAAAATAAAAAGCTGT
GAGGTTTTGTTTAATCCTTTTGATGACATCATTCCAAGGGAAATTAAAAGGCTGAAAAAAGA
GAAACCAGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAAATTTTAGTTTACTTT
CATTTGGAGAGGAAGCTGAGGAAGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGAAG
GGCAAAAGCAAAAGTAGTCATGACTTGCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGT
TGTAGAAAGTGAAAAAGGTGATGCACCAGATTTAGTTGATGATGGAGAAGATGAAAGTGCAG
AGCATGATGAATATATTGATGGTGATGAAAAGAACCTGATGAGAGAAAGAATTGCCAAAAAA
TTAAAAAAGGACACAAGTGCGAATGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATC
AGTCAGCCGCAGTGAAGAGCTCAGAAAAGAAGCAAGACAATTAAAACGGGAACTCTTAGCAG
CAAAACAAAAAAAGTAGAAAATGCAGCAAAACAAGCAGAAAAAAGAAGTGAAGAGGAAGAA
GCCCCTCCAGATGGTGCTGTTGCCGAATACAGAAGAGAAAAGCAAAAGTATGAAGCTTTGAG
GAAGCAACAGTCAAAGAAGGGAACTTCCCGGGAAGATCAGACCCTTGCACTGCTGAACCAGT
TTAAATCTAAACTCACTCAAGCAATTGCTGAAACACCTGAAAATGACATTCCTGAAACAGAA
GTAGAAGATGATGAAGGATGGATGTCACATGTACTTCAGTTTGAGGATAAAAGCAGAAAAGT
GAAAGATGCAAGCATGCAAGACTCAGATACATTTGAAATCTATGATCCTCGGAATCCAGTGA
ATAAAAGAAGGAGGGAAGAAAGCAAAAGCTGATGAGAGAGAAAAAAGAAAGAAGATAAAAT
GAGAATAATGATAACCAGAACTTGCTGGAAATGTGCCTACAATGGCCTTGTAACAGCCATTG
TTCCCAACAGCATCACTTAGGGGTGTGAAAAGAAGTATTTTTGAACCTGTTGTCTGGTTTTG
AAAAACAATTATCTTGTTTTGCAAATTGTGGAATGATGTAAGCAAATGCTTTTGGTTACTGG
TACATGTGTTTTTTCCTAGCTGACCTTTTATATTGCTAAATCTGAAATAAAATAACTTTCCT
TCCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 89

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50919

><subunit 1 of 1, 472 aa, 1 stop

><MW: 53847, pI: 5.75, NX(S/T): 2

MSNIYIQEPPTNGKVLLKTTAGDIDIELWSKEAPKACRNFIQLCLEAYYDNTIFHRVVPGFI

VQGGDPTGTGSGGESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFTLGRADELN

NKHTIFGKVTGDTVYNMLRLSEVDIDDDERPHNPHKIKSCEVLFNPFDDIIPREIKRLKKEK

PEEEVKKLKPKGTKNFSLLSFGEEAEEEEEVNRVSQSMKGKSKSSHDLLKDDPHLSSVPVV

ESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKKLKKDTSANVKSAGEGEVEKKSV

SRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEEAPPDGAVAEYRREKQKYEALRK

QQSKKGTSREDQTLALLNQFKSKLTQAIAETPENDIPETEVEDDEGWMSHVLQFEDKSRKVK

DASMQDSDTFEIYDPRNPVNKRRREESKKLMREKKERR

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 109-112 and 201-204

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.

amino acids 49-66

Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase

amino acids 96-140, 49-89 and 22-51

FIGURE 90

CGCCGCCGTTGGGGCTGGAAGTTCCCGCCAGGTCCGTGCCGGGCGAGAGAGATGCTGCCCGG
CCCGCCTCGGCTTTGAGGCGAGAGAAGTGTCCCAGACCCATTTCGCCTTGCTGACGGCGTCG
AGCCCTGGCCAGAC**ATG**TCCACAGGGTTCTCCTTCGGGTCCGGGACTCTGGGCTCCACCACC
GTGGCCGCCGGCGGGACCAGCACAGGCGGCGTTTTCTCCTTCGGAACGGGAACGTCTAGCAA
CCCTTCTGTGGGGCTCAATTTTGGAAATCTTGGAAGTACTTCAACTCCAGCAACTACATCTG
CTCCTTCAAGTGGTTTTGGAACCGGGCTCTTTGGATCTAAACCTGCCACTGGGTTCACTCTA
GGAGGAACAAATACAGGTGCCTTGCACACCAAGAGGCCTCAAGTGGTCACCAAATATGGAAC
CCTGCAAGGAAAACAGATGCATGTGGGGAAGACACCCATCCAAGTCTTTTTAGGAGTCCCCT
TCTCCAGACCTCCTCTAGGTATCCTCAGGTTTGCACCTCCAGAACCCCCGGAGCCCTGGAAA
GGAATCAGAGATGCTACCACCTACCCGCCTGGATGGAGTCTCGCTCTGTCGCCAGGCTGGAG
TGCAGTGGCACGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGAGTCTCCTGC
CTCAGCCTCTGAGTGTCTGGGGCTACAGGTGCCTGCAGGAGTCCTGGGGCCAGCTGGCCTCG
ATGTACGTCAGCACGCGGGAACGGTACAAGTGGCTGCGCTTCAGCGAGGACTGTCTGTACCT
GAACGTGTACGCGCCGGCGCGCGCGCCCGGGGATCCCCAGCTGCCAGTGATGGTCTGGTTCC
CGGGAGGCGCCTTCATCGTGGGCGCTGCTTCTTCGTACGAGGGCTCTGACTTGGCCGCCCGC
GAGAAAGTGGTGCTGGTGTTTCTGCAGCACAGGCTCGGCATCTTCGGCTTCCTGAGCACGGA
CGACAGCCACGCGCGCGGGAACTGGGGGCTGCTGGACCAGATGGCGGCTCTGCGCTGGGTGC
AGGAGAACATCGCAGCCTTCGGGGGAGACCCAGGAAATGTGACCCTGTTCGGCCAGTCGGCG
GGGGCCATGAGCATCTCAGGACTGATGATGTCACCCCTAGCCTCGGGTCTCTTCCATCGGGC
CATTTCCCAGAGTGGCACCGCGTTATTCAGACTTTTCATCACTAGTAACCCACTGAAAGTGG
CCAAGAAGGTTGCCCACCTGGCTGGATGCAACCACAACAGCACACAGATCCTGGTAAACTGC
CTGAGGGCACTATCAGGGACCAAGGTGATGCGTGTGTCCAACAAGATGAGATTCCTCCAACT
GAACTTCCAGAGAGACCCGGAAGAGATTATCTGGTCCATGAGCCCTGTGGTGGATGGTGTGG
TGATCCCAGATGACCCTTTGGTGCTCCTGACCCAGGGGAAGGTTTCATCTGTGCCCTACCTT
CTAGGTGTCAACAACCTGGAATTCAATTGGCTCTTGCCTTATAATATCACCAAGGAGCAGGT
ACCACTTGTGGTGGAGGAGTACCTGGACAATGTCAATGAGCATGACTGGAAGATGCTACGAA
ACCGTATGATGGACATAGTTCAAGATGCCACTTTCGTGTATGCCACACTGCAGACTGCTCAC
TACCACCGAGAAACCCCAATGATGGGAATCTGCCCTGCTGGCCACGCTACAACAAGGATGAA
AAGTACCTGCAGCTGGATTTTACCACAAGAGTGGGCA**TGA**AGCTCAAGGAGAAGAAGATGGC
TTTTTGGATGAGTCTGTACCAGTCTCAAAGACCTGAGAAGCAGAGGCAATTCTAAGGGTGGC
TATGCAGGAAGGAGCCAAAGAGGGGTTTGCCCCCACCATCCAGGCCCTGGGGAGACTAGCCA
TGGACATACCTGGGGACAAGAGTTCTACCCACCCCAGTTTAGAACTGCAGGAGCTCCCTGCT
GCCTCCAGGCCAAAGCTAGAGCTTTTGCCTGTTGTGTGGGACCTGCACTGCCCTTTCCAGCC
TGACATCCATGATGCCCCTCTACTTCACTGTTGACATCCAGTTAGGCCAGGCCCTGTCAAC
ACCACACTGTGCTCAGCTCTCCAGCCTCAGGACAACCTCTTTTTTTCCCTTCTTCAAATCCT
CCCACCCTTCAATGTCTCCTTGTGACTCCTTCTTATGGGAGGTCGACCCAGACTGCCACTGC
CCCTGTCACTGCACCCAGCTTGGCATTTACCATCCATCCTGCTCAACCTTGTTCCTGTCTGT
TCACATTGGCCTGGAGGCCTAGGGCAGGTTGTGACATGGAGCAAACTTTTGGTAGTTTGGGA
TCTTCTCTCCCACCCACACTTATCTCCCCCAGGGCCACTCCAAAGTCTATACACAGGGGTGG
TCTCTTCAATAAAGAAGTGTTGATTAGAAAAAAAAAA

FIGURE 91

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179

<subunit 1 of 1, 545 aa, 1 stop

<MW: 58934, pI: 9.45, NX(S/T): 4

MSTGFSFGSGTLGSTTVAAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATTSAPSSG

FGTGLFGSKPATGFTLGGTNTGALHTKRPQVVTKYGTLQGKQMHVGKTPIQVFLGVPFSRPP

LGILRFAPPEPPEPWKGIRDATTYPPGWSLALSPGWSAVARSRLTATSASRVQASLLPQPLS

VWGYRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLNVYAPARAPGDPQLPVMVWFPGGAF

IVGAASSYEGSDLAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIA

AFGGDPGNVTLFGQSAGAMSISGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVA

HLAGCNHNSTQILVNCLRALSGTKVMRVSNKMRFLQLNFQRDPEEIIWSMSPVVDGVVIPDD

PLVLLTQGKVSSVPYLLGVNNLEFNWLLPYNITKEQVPLVVEEYLDNVNEHDWKMLRNRMMD

IVQDATFVYATLQTAHYHRETPMMGICPAGHATTRMKSTCSWILPQEWA

Important features:

Signal peptide:

amino acids 1-29

Carboxylesterases type-B serine active site.

amino acids 312-327

Carboxylesterases type-B signature 2.

amino acids 218-228

N-glycosylation sites.

amino acids 318-321, 380-383 and 465-468

FIGURE 92

GAGAACAGGCCTGTCTCAGGCAGGCCCTGCGCCTCCTATGCGGAG**ATG**CTACTGCCACTGCT
GCTGTCCTCGCTGCTGGGCGGGTCCCAGGCTATGGATGGGAGATTCTGGATACGAGTGCAGG
AGTCAGTGATGGTGCCGGAGGGCCTGTGCATCTCTGTGCCCTGCTCTTTCTCCTACCCCCGA
CAAGACTGGACAGGGTCTACCCCAGCTTATGGCTACTGGTTCAAAGCAGTGACTGAGACAAC
CAAGGGTGCTCCTGTGGCCACAAACCACCAGAGTCGAGAGGTGGAAATGAGCACCCGGGGCC
GATTCCAGCTCACTGGGGATCCCGCCAAGGGGAACTGCTCCTTGGTGATCAGAGACGCGCAG
ATGCAGGATGAGTCACAGTACTTCTTTCGGGTGGAGAGAGGAAGCTATGTGACATATAATTT
CATGAACGATGGGTTCTTTCTAAAAGTAACAGTGCTCAGCTTCACGCCCAGACCCCAGGACC
ACAACACCGACCTCACCTGCCATGTGGACTTCTCCAGAAAGGGTGTGAGCGCACAGAGGACC
GTCCGACTCCGTGTGGCCTATGCCCCAGAGACCTTGTTATCAGCATTTCACGTGACAACAC
GCCAGCCCTGGAGCCCCAGCCCCAGGGAAATGTCCCATACCTGGAAGCCCAAAAAGGCCAGT
TCCTGCGGCTCCTCTGTGCTGCTGACAGCCAGCCCCCTGCCACACTGAGCTGGGTCCTGCAG
AACAGAGTCCTCTCCTCGTCCCATCCCTGGGGCCCTAGACCCCTGGGGCTGGAGCTGCCCGG
GGTGAAGGCTGGGGATTCAGGGCGCTACACCTGCCGAGCGGAGAACAGGCTTGGCTCCCAGC
AGCGAGCCCTGGACCTCTCTGTGCAGTATCCTCCAGAGAACCTGAGAGTGATGGTTTCCCAA
GCAAACAGGACAGTCCTGGAAAACCTTGGGAACGGCACGTCTCTCCCAGTACTGGAGGGCCA
AAGCCTGTGCCTGGTCTGTGTCACACACAGCAGCCCCCAGCCAGGCTGAGCTGGACCCAGA
GGGGACAGGTTCTGAGCCCCTCCCAGCCCTCAGACCCCGGGGTCCTGGAGCTGCCTCGGGTT
CAAGTGGAGCACGAAGGAGAGTTCACCTGCCACGCTCGGCACCCACTGGGCTCCCAGCACGT
CTCTCTCAGCCTCTCCGTGCACTATAAGAAGGGACTCATCTCAACGGCATTCTCCAACGGAG
CGTTTCTGGGAATCGGCATCACGGCTCTTCTTTTCCTCTGCCTGGCCCTGATCATCATGAAG
ATTCTACCGAAGAGACGGACTCAGACAGAAACCCCGAGGCCCAGGTTCTCCCGGCACAGCAC
GATCCTGGATTACATCAATGTGGTCCCGACGGCTGGCCCCCTGGCTCAGAAGCGGAATCAGA
AAGCCACACCAAACAGTCCTCGGACCCCTCCTCCACCAGGTGCTCCCTCCCCAGAATCAAAG
AAGAACCAGAAAAAGCAGTATCAGTTGCCCAGTTTCCCAGAACCCAAATCATCCACTCAAGC
CCCAGAATCCCAGGAGAGCCAAGAGGAGCTCCATTATGCCACGCTCAACTTCCCAGGCGTCA
GACCCAGGCCTGAGGCCCGGATGCCCAAGGGCACCCAGGCGGATTATGCAGAAGTCAAGTTC
CAA**TGA**GGGTCTCTTAGGCTTTAGGACTGGGACTTCGGCTAGGGAGGAAGGTAGAGTAAGAG
GTTGAAGATAACAGAGTGCAAAGTTTCCTTCTCTCCCTCTCTCTCTCTCTTTCTCTCTCTCT
CTCTCTTTCTCTCTCTTTTAAAAAAACATCTGGCCAGGGCACAGTGGCTCACGCCTGTAATC
CCAGCACTTTGGGAGGTTGAGGTGGGCAGATCGCCTGAGGTCGGGAGTTCGAGACCAGCCTG
GCCAACTTGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCATGGTGGCAGG
CGCCTGTAATCCTACCTACTTGGGAAGCTGAGGCAGGAGAATCACTTGAACCTGGGAGACGG
AGGTTGCAGTGAGCCAAGATCACACCATTGCACGCCAGCCTGGGCAACAAAGCGAGACTCCA
TCTCAAAAAAAAATCCTCCAAATGGGTTGGGTGTCTGTAATCCCAGCACTTTGGGAGGCTA
AGGTGGGTGGATTGCTTGAGCCCAGGAGTTCGAGACCAGCCTGGGCAACATGGTGAAACCCC
ATCTCTACAAAAAATACAAAACATAGCTGGGCTTGGTGGTGTGTGCCTGTAGTCCCAGCTGT
CAGACATTTAAACCAGAGCAACTCCATCTGGAATAGGAGCTGAATAAAATGAGGCTGAGACC
TACTGGGCTGCATTCTCAGACAGTGGAGGCATTCTAAGTCACAGGATGAGACAGGAGGTCCG
TACAAGATACAGGTCATAAAGACTTTGCTGATAAAACAGATTGCAGTAAAGAAGCCAACCAA
ATCCCACCAAAACCAAGTTGGCCACGAGAGTGACCTCTGGTCGTCCTCACTGCTACACTCCT
GACAGCACCATGACAGTTTACAAATGCCATGGCAACATCAGGAAGTTACCCGATATGTCCCA
AAAGGGGGAGGAATGAATAATCCACCCCTTGTTTAGCAAATAAGCAAGAAATAACCATAAAA
GTGGGCAACCAGCAGCTCTAGGCGCTGCTCTTGTCTATGGAGTAGCCATTCTTTTGTTCCTT
TACTTTCTTAATAAACTTGCTTTCACCTTAAAAAAA

FIGURE 93

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002
><subunit 1 of 1, 544 aa, 1 stop
><MW: 60268, pI: 9.53, NX(S/T): 3

MLLPLLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSFSYPRQDWTGSTPAYGYWFK
AVTETTKGAPVATNHQSREVEMSTRGRFQLTGDPAKGNCSLVIRDAQMQDESQYFFRVERGS
YVTYNFMNDGFFLKVTVLSFTPRPQDHNTDLTCHVDFSRKGVSAQRTVRLRVAYAPRDLVIS
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLLCAADSQPPATLSWVLQNRVLSSSHPWGPRPL
GLELPGVKAGDSGRYTCRAENRLGSQQRALDLSVQYPPENLRVMVSQANRTVLENLGNGTSL
PVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSPSQPSDPGVLELPRVQVEHEGEFTCHARHP
LGSQHVSLSLSVHYKKGLISTAFSNGAFLGIGITALLFLCLALIIMKILPKRRTQTETPRPR
FSRHSTILDYINVVPTAGPLAQKRNQKATPNSPRTPPPPGAPSPESKKNQKKQYQLPSFPEP
KSSTQAPESQESQEELHYATLNFPGVRPRPEARMPKGTQADYAEVKFQ

Important features:
Signal peptide:
amino acids 1-15

Transmembrane domain:
amino acids 399-418

N-glycosylation site.
amino acids 100-103, 297-300 and 306-309

Immunoglobulins and major histocompatibility complex proteins
signature.
amino acids 365-371

FIGURE 94

TGAAGAGTAATAGTTGGAATCAAAAGAGTCAACGCA**ATG**AACTGTTATTTACTGCTGCGTTT
TATGTTGGGAATTCCTCTCCTATGGCCTTGTCTTGGAGCAACAGAAAACTCTCAAACAAAGA
AAGTCAAGCAGCCAGTGCGATCTCATTTGAGAGTGAAGCGTGGCTGGGTGTGGAACCAATTT
TTTGTACCAGAGGAAATGAATACGACTAGTCATCACATCGGCCAGCTAAGATCTGATTTAGA
CAATGGAAACAATTCTTTCCAGTACAAGCTTTTGGGAGCTGGAGCTGGAAGTACTTTTATCA
TTGATGAAAGAACAGGTGACATATATGCCATACAGAAGCTTGATAGAGAGGAGCGATCCCTC
TACATCTTAAGAGCCCAGGTAATAGACATCGCTACTGGAAGGGCTGTGGAACCTGAGTCTGA
GTTTGTCATCAAAGTTTCGGATATCAATGACAATGAACCAAAATTCCTAGATGAACCTTATG
AGGCCATTGTACCAGAGATGTCTCCAGAAGGAACATTAGTTATCCAGGTGACAGCAAGTGAT
GCTGACGATCCCTCAAGTGGTAATAATGCTCGTCTCCTCTACAGCTTACTTCAAGGCCAGCC
ATATTTTTCTGTTGAACCAACAACAGGAGTCATAAGAATATCTTCTAAAATGGATAGAGAAC
TGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGTCAGCCAGGAGCGTTG
TCTGGAACAACAAGTGTATTAATTAAACTTTCAGATGTTAATGACAATAAGCCTATATTTAA
AGAAAGTTTATACCGCTTGACTGTCTCTGAATCTGCACCCACTGGGACTTCTATAGGAACAA
TCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTACAGCATTGAAGAGGAT
GATTCGCAAACATTTGACATTATTACTAATCATGAAACTCAAGAAGGAATAGTTATATTAAA
AAAGAAAGTGGATTTTGAGCACCAGAACCACTACGGTATTAGAGCAAAAGTTAAAAACCATC
ATGTTCCTGAGCAGCTCATGAAGTACCACACTGAGGCTTCCACCACTTTCATTAAGATCCAG
GTGGAAGATGTTGATGAGCCTCCTCTTTTCCTCCTTCCATATTATGTATTTGAAGTTTTTGA
AGAAACCCCACAGGGATCATTTGTAGGCGTGGTGTCTGCCACAGACCCAGACAATAGGAAAT
CTCCTATCAGGTATTCTATTACTAGGAGCAAAGTGTTCAATATCAATGATAATGGTACAATC
ACTACAAGTAACTCACTGGATCGTGAAATCAGTGCTTGGTACAACCTAAGTATTACAGCCAC
AGAAAAATACAATATAGAACAGATCTCTTCGATCCCACTGTATGTGCAAGTTCTTAACATCA
ATGATCATGCTCCTGAGTTCTCTCAATACTATGAGACTTATGTTTGTGAAAATGCAGGCTCT
GGTCAGGTAATTCAGACTATCAGTGCAGTGGATAGAGATGAATCCATAGAAGAGCACCATTT
TTACTTTAATCTATCTGTAGAAGACACTAACAATTCAAGTTTTACAATCATAGATAATCAAG
ATAACACAGCTGTCATTTTGACTAATAGAACTGGTTTTAACCTTCAAGAAGAACCTGTCTTC
TACATCTCCATCTTAATTGCCGACAATGGAATCCCGTCACTTACAAGTACAAACACCCTTAC
CATCCATGTCTGTGACTGTGGTGACAGTGGGAGCACACAGACCTGCCAGTACCAGGAGCTTG
TGCTTTCCATGGGATTCAAGACAGAAGTTATCATTGCTATTCTCATTTGCATTATGATCATA
TTTGGGTTTATTTTTTTGACTTTGGGTTTAAAACAACGGAGAAAACAGATTCTATTTCCTGA
GAAAAGTGAAGATTTCAGAGAGAATATATTCCAATATGATGATGAAGGGGGTGGAGAAGAAG
ATACAGAGGCCTTTGATATAGCAGAGCTGAGGAGTAGTACCATAATGCGGGAACGCAAGACT
CGGAAAACCACAAGCGCTGAGATCAGGAGCCTATACAGGCAGTCTTTGCAAGTTGGCCCCGA
CAGTGCCATATTCAGGAAATTCATTCTGGAAAAGCTCGAAGAAGCTAATACTGATCCGTGTG
CCCCTCCTTTTGATTCCCTCCAGACCTACGCTTTTGAGGGAACAGGGTCATTAGCTGGATCC
CTGAGCTCCTTAGAATCAGCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTAATGAGTT
GGGACCTCGCTTTAAAAGATTAGCATGCATGTTTGGTTCTGCAGTGCAGTCAAATAAT**TAG**G
GCTTTTTACCATCAAAATTTTTAAAAGTGCTAATGTGTATTCGAACCCAATGGTAGTCTTAA
AGAGTTTTGTGCCCTGGCTCTATGGCGGGGAAAGCCCTAGTCTATGGAGTTTTCTGATTTCC
CTGGAGTAAATACTCCATGGTTATTTTAAGCTACCTACATGCTGTCATTGAACAGAGATGTG
GGGAGAAATGTAAACAATCAGCTCACAGGCATCAATACAACCAGATTTGAAGTAAAATAATG
TAGGAAGATATTAAAAGTAGATGAGAGGACACAAGATGTAGTCGATCCTTATGCGATTATAT
CATTATTTACTTAGGAAAGAGTAAAAATACCAAACGAGAAATTTAAAGGAGCAAAAATTTG
CAAGTCAAATAGAAATGTACAAATCGAGATAACATTTACATTTCTATCATATTGACATGAAA
ATTGAAAATGTATAGTCAGAGAAATTTTCATGAATTATTCCATGAAGTATTGTTTCCTTTAT
TTAAA

FIGURE 95

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906

><subunit 1 of 1, 772 aa, 1 stop

><MW: 87002, pI: 4.64, NX(S/T): 8

MNCYLLLRFMLGIPLLWPCLGATENSQTKKVKQPVRSHLRVKRGWVWNQFFVPEEMNTTSHH

IGQLRSDLDNGNNSFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT

GRAVEPESEFVIKVSDINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARL

LYSLLQGQPYFSVEPTTGVIRISSKMDRELQDEYWVIIQAKDMIGQPGALSGTTSVLIKLSD

VNDNKPIFKESLYRLTVSESAPTGTSIGTIMAYDNDIGENAEMDYSIEEDDSQTFDIITNHE

TQEGIVILKKKVDFEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLL

PYYVFEVFEETPQGSFVGVVSATDPDNRKSPIRYSITRSKVFNINDNGTITTSNSLDREISA

WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDR

DESIEEHHFYFNLSVEDTNNSSFTIIDNQDNTAVILTNRTGFNLQEEPVFYISILIADNGIP

SLTSTNTLTIHVCDCGDSGSTQTCQYQELVLSMGFKTEVIIAILICIMIIFGFIFLTLGLKQ

RRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTSAEIRSLY

RQSLQVGPDSAIFRKFILEKLEEANTDPCAPPFDSLQTYAFEGTGSLAGSLSSLESAVSDQD

ESYDYLNELGPRFKRLACMFGSAVQSNN

Important features:
Signal peptide:
amino acids 1-21

Transmembrane domain:
amino acids 597-617

N-glycosylation sites.
amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518, 516-519 and 534-537

Cadherins extracellular repeated domain signature.
amino acids 136-146 and 244-254

FIGURE 96

ATTTCAAGGCCAGCCATATTTTTNTGTTGAACCAACAACAGGAGTCATAAGAATATTTTNTA

AAATGGATAGAGAACTGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGT

CAGCCAGGAGCGTTGTNTGGAACAACAAGTGTATTAATTAAACTTTCAGATGTTAATGACAA

TAAGCCTATATTTAAAGAAAGTTTATACCGCTTGACTGTNTNTGAATCTGCACCCACTGGGA

NTTNTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTAC

AGCATTGAAGAGGATGATTCGCAAACATTTGACATTATT

FIGURE 97

GCAACCTCAGCTTCTAGTATCCAGACTCCAGCGCCGCCCCGGGCGCGGACCCCAACCCCGAC
CCAGAGCTTCTCCAGCGGCGGCGCAGCGAGCAGGGCTCCCCGCCTTAACTTCCTCCGCGGGG
CCCAGCCACCTTCGGGAGTCCGGGTTGCCCACCTGCAAACTCTCCGCCTTCTGCACCTGCCA
CCCCTGAGCCAGCGCGGGCCCCCGAGCGAGTC**ATG**GCCAACGCGGGGCTGCAGCTGTTGGGC
TTCATTCTCGCCTTCCTGGGATGGATCGGCGCCATCGTCAGCACTGCCCTGCCCCAGTGGAG
GATTTACTCCTATGCCGGCGACAACATCGTGACCGCCCAGGCCATGTACGAGGGGCTGTGGA
TGTCCTGCGTGTCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAAT
CTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGAT
AGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGC
AGAAGATGAGGATGGCTGTCATTGGGGGTGCGATATTTCTTCTTGCAGGTCTGGCTATTTTA
GTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCCCAGT
CAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCC
TTCTGGGAGGTGCCCTACTTTGCTGTTCCTGTCCCCGAAAAACAACCTCTTACCCAACACCA
AGGCCCTATCCAAAACCTGCACCTTCCAGCGGGAAAGACTACGTG**TGA**CACAGAGGCAAAAG
GAGAAAATCATGTTGAAACAAACCGAAATGGACATTGAGATACTATCATTAACATTAGGAC
CTTAGAATTTTGGGTATTGTAATCTGAAGTATGGTATTACAAAACAAACAAACAAACAAAAA
ACCCATGTGTTAAAATACTCAGTGCTAAACATGGCTTAATCTTATTTTATCTTCTTTCCTCA
ATATAGGAGGGAAGATTTTTCCATTTGTATTACTGCTTCCCATTGAGTAATCATACTCAAAT
GGGGGAAGGGGTGCTCCTTAAATATATATAGATATGTATATATACATGTTTTTCTATTAAAA
ATAGACAGTAAAATACTATTCTCATTATGTTGATACTAGCATACTTAAAATATCTCTAAAT
AGGTAAATGTATTTAATTCCATATTGATGAAGATGTTTATTGGTATATTTTCTTTTTCGTCC
TTATATACATATGTAACAGTCAAATATCATTTACTCTTCTTCATTAGCTTTGGGTGCCTTTG
CCACAAGACCTAGCCTAATTTACCAAGGATGAATTCTTTCAATTCTTCATGCGTGCCCTTTT
CATATACTTATTTTATTTTTTACCATAATCTTATAGCACTTGCATCGTTATTAAGCCCTTAT
TTGTTTTGTGTTTCATTGGTCTCTATCTCCTGAATCTAACACATTTCATAGCCTACATTTTA
GTTTCTAAAGCCAAGAAGAATTTATTACAAATCAGAACTTTGGAGGCAAATCTTTCTGCATG
ACCAAAGTGATAAATTCCTGTTGACCTTCCCACACAATCCCTGTACTCTGACCCATAGCACT
CTTGTTTGCTTTGAAAATATTTGTCCAATTGAGTAGCTGCATGCTGTTCCCCCAGGTGTTGT
AACACAACTTTATTGATTGAATTTTTAAGCTACTTATTCATAGTTTTATATCCCCCTAAACT
ACCTTTTTGTTCCCCATTCCTTAATTGTATTGTTTTCCCAAGTGTAATTATCATGCGTTTTA
TATCTTCCTAATAAGGTGTGGTCTGTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATA
ATCTGGTGACAAATATTCTCTCTGTAGCTGTAAGCAAGTCACTTAATCTTTCTACCTCTTTT
TTCTATCTGCCAAATTGAGATAATGATACTTAACCAGTTAGAAGAGGTAGTGTGAATATTAA
TTAGTTTATATTACTCTTATTCTTTGAACATGAACTATGCCTATGTAGTGTCTTTATTTGCT
CAGCTGGCTGAGACACTGAAGAAGTCACTGAACAAAACCTACACACGTACCTTCATGTGATT
CACTGCCTTCCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACACACATACCTTCAT
GTGGTTCAGTGCCTTCCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACGCACATAC
CTTCATGTGGCTCAGTGCCTTCCTCTCTACCAGTCTATTTCCATTCTTTCAGCTGTGTCT
GACATGTTTGTGCTCTGTTCCATTTTAACAACTGCTCTTACTTTTCCAGTCTGTACAGAATG
CTATTTCACTTGAGCAAGATGATGTAATGGAAAGGGTGTTGGCACTGGTGTCTGGAGACCTG
GATTTGAGTCTTGGTGCTATCAATCACCGTCTGTGTTTGAGCAAGGCATTTGGCTGCTGTAA
GCTTATTGCTTCATCTGTAAGCGGTGGTTTGTAATTCCTGATCTTCCCACCTCACAGTGATG
TTGTGGGGATCCAGTGAGATAGAATACATGTAAGTGTGGTTTTGTAATTTAAAAAGTGCTAT
ACTAAGGGAAAGAATTGAGGAATTAACTGCATACGTTTTGGTGTTGCTTTTCAAATGTTTGA
AAATAAAAAAAATGTTAAG

FIGURE 98

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185

><subunit 1 of 1, 211 aa, 1 stop

><MW: 22744, pI: 8.51, NX(S/T): 1

MANAGLQLLGFILAFLGWIGAIVSTALPQWRIYSYAGDNIVTAQAMYEGLWMSCVSQSTGQI
QCKVFDSLLNLSSTLQATRALMVVGILLGVIAIFVATVGMKCMKCLEDDEVQKMRMAVIGGA
IFLLAGLAILVATAWYGNRIVQEFYDPMTPVNARYEFGQALFTGWAAASLCLLGGALLCCSC
PRKTTSYPTPRPYPKPAPSSGKDYV

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-102, 118-142 and 161-187

N-glycosylation site.

amino acids 72-75

PMP-22 / EMP / MP20 family proteins

amino acids 70-111

ABC-2 type transport system integral membrane protein

amino acids 119-133

FIGURE 99

TTCTGGCCAAACCCGGGGCTNCAGCTGTTGGGCTTCATCTCGCCTTCCTGGGATGGATCGGC

GCCATCNTCACACTGCCCTTCCCCAGTGGAGGATTTTACTCCCTATGCTGGCGACAACATCG

TGACCGCCCAGCCCATGTACGAGGGGCTGTGGATGTCCNGCGTGTCGCAGAGCACCGGGCAG

ATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCCGTGC

CTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGA

AGTGTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGC

GCGATATTTCTTCTTGCAGGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAAN

CNTTCAACANTTCTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCA

GGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCT

GTTCCTGTCCC

FIGURE 100

ACCCTTGACCCAACGCGGCCCCCCGACCGNTTCATGGCCAAACGCGGGNCTCCAGCTGTTGG

GCTTCATTCTCCCCTTCCTGGGATGGACCGGCGCCCATCNTCAGCACTGCCCTGCCCCAGTG

GAGGATTTACTCCTATNCCGGCNACAACATCGTGACCGCCCAGGCCNTGTACGAGGGGCTGT

GGATGTCCTGCGTGTCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCCTTGCT

GAATCTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAG

TGATAGCAATCTTNNTGGCCACCGTTGTNNNTGAAGTGTATGAAGTGCTTGGAAGACGATGA

GGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAGGTCTGGCTA

TTTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCGA

FIGURE 101

GGGCCCGACCATTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCGCGC

CATCNTCAGACTCCCTGCCCCATGGAGATTTNNCCTATGCTGGCGACAACATCNTGACCCCC

AGCCATGTACGAGGGGCTTTGAACGTCNGCGTGTCGCAGANCACCGGGCAGATCCAGTGCAA

AGTCTTTGACTCCTTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT

TGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGT

GCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTT

CTTGCAGGTCTGGCTATTTNNNGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAAT

TCTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGC

TGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCCTGCGA

FIGURE 102

ATTCTCCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTTCCCCCANTGGAGGATTNAC

TCCTATGCTGGCGACAACATCGTGACCCCCCAGGCCATTTACCGAGGGGCTTTGGATGTCNT

GCNTGTCGCAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAG

CAGCACATTGCAAGCAACCCGTGCCTTGATGGGGTTGGCATCCTCCTGGGAGTGATAGCAAC

CTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCCAGAAG

ATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTGTTGCAGGTCTGGCTATTTTAGTNGC

CACAGCATGGTATGGCAATAGANTNNTTCNNGNNNTCTATGACCCTATGACCCCAGTCAATG

CCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTG

GGAGGTGCCCTACTTTGCTGTTCCTGTCCC

FIGURE 103

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTTGACCCTTGCTGAATCTGAGCAGCACATTNC

AAGCAACCCCTTGCCTTGAAGGTGGTTGNCATCCCCCCTGGGAGTGAATAGCAATCTTTGTG

GCCACCGTTGGCATGAAGTNTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT

GGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAGGTCTGGCTATTTTAGTNNCCACAGCAT

GGTATGGCAATAGNATNNTTCGNGGNTTCTATGACCCTATGACCCCAGTCAATGCCAGGTAC

GAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGC

CCTACTTTGCTGTTCCTGTCCCCGAA

FIGURE 104

AGCAATGCCCTGCCCCCAGTGGAGGATTAATTCCTATGNTGGGGACAACATTGTGACNGCCC
AGGCCATGTACGGGGGGCTGTGGATGTCCTGCGTGTCGCAGAGCACCGGGCAGATCCAGTGC
AAAGTNTTTGACTCCTTGCTGAATTTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGT
GGTTGGCATCTTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTGGNAATGAAGTGTATGA
AGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTT
CTTNTTGCAGGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAATNGTTCAAGA
ATTTTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTTTNTTCACTG
GCTGGGCTGCTGCTTNTTTCTGCCTTNTGGGAGGTGCCCTANTTTGCTGTTCCTGCGAACC

FIGURE 105

TCATAGGGGGGCGCGATATTTTTTCTTGCAGGTNTGGTTATTTTAGTTGCCACAGCATGGTA

TGGCAATAGAATCGTTCAAGAATTNTATGACCCTATGACCCCAGTCAATGCCAGGTACGAAT

TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTTNTGGGAGGTGCCCTA

CTTTGCTGTTCCTG

FIGURE 106

TTCCTGGGATGGATCCGCCCCCATCNTCACATGCCCTGCCCCNTGGAGATTTACNCCTATGC

TGGCGAACAACATCNTGACCGCCCAGGCCATGTACGAGGGGCTGTGGAATGTCCTGCGTGTC

CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACAT

TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGG

CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT

GGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAGGTCTGGCTATTTTAGNNGCCACAGCAT

GGTATGGCAATCAGACCCNNTCANAAACTCTATGACCCTATGACCCCAGTCAATGCCAGGTA

CGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTG

CCCTACTTTGCTGTTCCTGTCCCCGAAAAACAACCTCTTACCCACG

FIGURE 107

CGGGGCTGCAGCTGTTGGGCTTCATCTCGCTTCCTGGGATGGAATCGGCGCCATCGTCAGCA

CTGCCCTGCCCCATGGAGGATTTACTCNTATGCTGGCGACAACATCGTGACCNCCCAGGCCA

TGTACGAGGGGCTGTGGATGTCNGCGTGTCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCT

TTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCNTGCCTTGATGGTGGTTGGCA

TCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTG

GAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGC

AGGTCTGGCTATTTNTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTAT

GACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGC

TGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCCTGCGAA

FIGURE 108

GCGTGCCGTCAGCTCGCCGGGCACCGCGGCCTCGCCCTCGCCCTCCGCCCCTGCGCCTGCAC
CGCGTAGACCGACCCCCCCCTCCAGCGCGCCCACCCGGTAGAGGACCCCCGCCCGTGCCCCG
ACCGGTCCCCGCCTTTTTGTAAAACTTAAAGCGGGCGCAGCATTAACGCTTCCCGCCCCGGT
GACCTCTCAGGGGTCTCCCCGCCAAAGGTGCTCCGCCGCTAAGGAAC**ATGG**CGAAGGTGGAG
CAGGTCCTGAGCCTCGAGCCGCAGCACGAGCTCAAATTCCGAGGTCCCTTCACCGATGTTGT
CACCACCAACCTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTTTAAGGTGAAGACTA
CAGCACCACGTAGGTACTGTGTGAGGCCCAACAGCGGAATCATCGATGCAGGGGCCTCAATT
AATGTATCTGTGATGTTACAGCCTTTCGATTATGATCCCAATGAGAAAAGTAAACACAAGTT
TATGGTTCAGTCTATGTTTGCTCCAACTGACACTTCAGATATGGAAGCAGTATGGAAGGAGG
CAAAACCGGAAGACCTTATGGATTCAAAACTTAGATGTGTGTTTGAATTGCCAGCAGAGAAT
GATAAACCACATGATGTAGAAATAAATAAAATTATATCCACAACTGCATCAAAGACAGAAAC
ACCAATAGTGTCTAAGTCTCTGAGTTCTTCTTTGGATGACACCGAAGTTAAGAAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTCAGAGGCTACGGGAGGAGAACAAGCAGTTCAAG
GAAGAAGATGGACTGCGGATGAGGAAGACAGTGCAGAGCAACAGCCCCATTTCAGCATTAGC
CCCAACTGGGAAGGAAGAAGGCCTTAGCACCCGGCTCTTGGCTCTGGTGGTTTTGTTCTTTA
TCGTTGGTGTAATTATTGGGAAGATTGCCTTG**TAG**AGGTAGCATGCACAGGATGGTAAATTG
GATTGGTGGATCCACCATATCATGGGATTTAAATTTATCATAACCATGTGTAAAAGAAATT
AATGTATGATGACATCTCACAGGTCTTGCCTTTAAATTACCCCTCCCTGCACACACATACAC
AGATACACACACACAAATATAATGTAACGATCTTTTAGAAAGTTAAAAATGTATAGTAACTG
ATTGAGGGGGAAAAAGAATGATCTTTATTAATGACAAGGGAAACCATGAGTAATGCCACAAT
GGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGCTGGATTACCTC
TCTTAAAATGACACCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTGGAGCCCAGCAT
GCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCCAGGCTG
CTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGA
AGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGT
TGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAA
GCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACTGTTATTCAGAGATGTTTAATGCATA
TTTAACTTATTTAATGTATTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGC
TGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTGTGGGCTCCTCT
GTCTCTGGAGAGTCTGGTCATGTGGAGGTGGGGTTTATTGGGATGCTGGAGAAGAGCTGCCA
GGAAGTGTTTTTTCTGGGTCAGTAAATAACAACTGTCATAGGGAGGGAAATTCTCAGTAGTG
ACAGTCAACTCTAGGTTACCTTTTTTAATGAAGAGTAGTCAGTCTTCTAGATTGTTCTTATA
CCACCTCTCAACCATTACTCACACTTCCAGCGCCCAGGTCCAAGTCTGAGCCTGACCTCCCC
TTGGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAAGCGAGGGC
ACCAGCAGTTGTGGGTGGGGAGCAAGGGAAGAGAGAAACTCTTCAGCGAATCCTTCTAGTAC
TAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATAAAAGACCAACCCAGTTCTGTTTGA
CTATGTAGCATCTTGAAAAGAAAAATTATAATAAAGCCCCAAAATTAAGAAAA

FIGURE 109

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977

<subunit 1 of 1, 243 aa, 1 stop

<MW: 27228, pI: 7.43, NX(S/T): 2

MAKVEQVLSLEPQHELKFRGPFTDVVTTNLKLGNPTDRNVCFKVKTTAPRRYCVRPNSGIID
AGASINVSVMLQPFDYDPNEKSKHKFMVQSMFAPTDTSDMEAVWKEAKPEDLMDSKLRCVFE
LPAENDKPHDVEINKIISTTASKTETPIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREE
NKQFKEEDGLRMRKTVQSNSPISALAPTGKEEGLSTRLLALVVLFFIVGVIIGKIAL

Important features:

Transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

FIGURE 110

GTCAGTCTTCTAGATTGTCCTTATCCCACCTTTCAACCANTACTCACATTTCNAGCGCCCAG

GTCCANGTCTGAGCCTGACTTCCCCTTGGGGACCTAGCCTGGAGTCAGGACAATGGNTCGGG

CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTTTGGGTGGGGAGCAAGGGNNGAGAGAAA

CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATA

AAAGACNAACCCAGTTCTGTTTGACTATGTAGCATCTTGAAAAGAAAAATTATAATAAAGCC

CCAAAATTAAGAATTCTTTTGTCATTTTGTCACATTTGCTCTATGGGGGGAATTATTATTTT

ATCATTTTTATTATTTTGCCATTGGAAGGTTAACTTTAAAATGAGC

FIGURE 111

TATTGTAAAGGCCATTTTAAACCATTGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT

AAATGACACCNTTCCTCGCCTGTTGGTGCTGGCCNTTGGGGAGCTGGAGCCCCAGCATGCTG

GGGAGTGCGGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCCAGGCTGCTTT

CCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAAGCC

CAAAGGAATTGCCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGTTGA

CTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCT

AAATTGTATTGGTTCATGTAGTGAAGTCAAACTGTTATTCAGAGATGTTTAATGCATATTTA

ACTTATTTAATGTATTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGCTGCG

TGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTG

FIGURE 112

CCCTGGTGGTTTTGTTCTTTAATTCGTTGGTGTAATTNTTGGGAAGATTGCTTGTAGAGGTA

GNATGCACCNGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTTAT

CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTTAAATT

ACCCATCCCTGNANACACATACACAGATACACANANACAAATNTAATGTAACGATNTTTTAG

AAAGTTAAAAATGTATAGTAAC

FIGURE 113

GGTGGCCCATTCCCGGCCCAGGCTGCTTTCCGGTNTTCAGTTCTGTCCAAGCCATCAGCTCC

TTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATNAGACGTAC

TTGTNATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGT

GCTTTGTTCANTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACTG

TTATTCAGAGATGTTTAATGCATATTTAANTTATTTAATGTATTTNATNTCATGTTTTCTTA

TTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAANTNTGTTGGGTGAACTGGTATTGC

TGCTGGAGGGCTGTGGGCTCCTCTGTCTTTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 114

TGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTTGATGAACAGAGTC

AGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTG

TGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGAC

CAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACTGTTATTCAGAGATGTTTAATGC

ATATTTAACTTATTTAATGTATTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAA

TGCTGCGTGC

FIGURE 115

AAACCTTTAAAAGTTGAGGGGAAAAGAATGATCCTTTATTAATGACAAGGGAAACCNTGNGT

AATGCCACAATGGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGC

TGGATTACCTCTCTTAAAATGACACCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTN

GAGCCCAGCATGCTGGGGAGTGCGGTCTGCTCCACACAGTAGTCCCCANGTGGCCCANTCCC

GGCCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGANTGATGA

ACAGAGTCAGAAGCCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA

GAGGCGTGTGTTGANTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCANTT

AAAGGGNCCAAGNTAAATTTGTATTGGTTCATGTAGTGAAGTCAAANTGTTATTCAGAGATG

TTTAATGCATATTTAANTTATTTAATGTATTTCATNTCATGTTTTCTTATTGTCACAAGGGT

ACAGTTAATGCTGCGTGCTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

FIGURE 116

GGCCCTTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC

CACGTGGCCCACTCCCGGCCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGC

TCCTTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACG

TACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGC

AGTGCTTTGTTCACTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAA

CTGTTATTCAGAGATGTTTAATGCATATTTAACTTATTTAATGTATTTCATCTCATGTTTTC

TTATTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTAT

TGCTGCTGGAGGGCTGTGGGCTCCTCTGTCTCTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 117

GCGAGCTCCGGGTGCTGTGGCCCGGCCTTGGCGGGGCGGCCTCCGGCTCAGGCTGGCTGAGA
GGCTCCCAGCTGCAGCGTCCCCGCCCGCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
CTCGGGGACCAAACAAGCCTGGCAGGGTCTCACTTTGTTGCCCAGGCTGGAGTTCAGTGCCA
TGATCATGGTTTACTGCAGCCTTGACCTCCTGGGTTCAAGCGATCCTGCTGAGTAGCTGGGA
CTACAGGACAAAATTAGAAGATCAAA**ATG**GAAAATATGCTGCTTTGGTTGATATTTTTCACC
CCTGGGTGGACCCTCATTGATGGATCTGAAATGGAATGGGATTTTATGTGGCACTTGAGAAA
GGTACCCCGGATTGTCAGTGAAAGGACTTTCCATCTCACCAGCCCCGCATTTGAGGCAGATG
CTAAGATGATGGTAAATACAGTGTGTGGCATCGAATGCCAGAAAGAACTCCCAACTCCCAGC
CTTTCTGAATTGGAGGATTATCTTTCCTATGAGACTGTCTTTGAGAATGGCACCCGAACCTT
AACCAGGGTGAAAGTTCAAGATTTGGTTCTTGAGCCGACTCAAAATATCACCACAAAGGGAG
TATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCGACAGCAGGTTCAGCATCTTGGACAAA
AGGTTCTTAACCAATTTCCCTTTCAGCACAGCTGTGAAGCTTTCCACGGGCTGTAGTGGCAT
TCTCATTTCCCCTCAGCATGTTCTAACTGCTGCCCACTGTGTTCATGATGGAAAGGACTATG
TCAAAGGGAGTAAAAAGCTAAGGGTAGGGTTGTTGAAGATGAGGAATAAAAGTGGAGGCAAG
AAACGTCGAGGTTCTAAGAGGAGCAGGAGAGAAGCTAGTGGTGGTGACCAAAGAGAGGGTAC
CAGAGAGCATCTGCAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAAAAAATCTGGCCGGGGTC
AGAGGATTGCCGAAGGGAGGCCTTCCTTTCAGTGGACCCGGGTCAAGAATACCCACATTCCG
AAGGGCTGGGCACGAGGAGGCATGGGGGACGCTACCTTGGACTATGACTATGCTCTTCTGGA
GCTGAAGCGTGCTCACAAAAAGAAATACATGGAACTTGGAATCAGCCCAACGATCAAGAAAA
TGCCTGGTGGAATGATCCACTTCTCAGGATTTGATAACGATAGGGCTGATCAGTTGGTCTAT
CGGTTTTGCAGTGTGTCCGACGAATCCAATGATCTCCTTTACCAATACTGCGATGCTGAGTC
GGGCTCCACCGGTTCGGGGGTCTATCTGCGTCTGAAAGATCCAGACAAAAAGAATTGGAAGC
GCAAATCATTGCGGTCTACTCAGGGCACCAGTGGGTGGATGTCCACGGGGTTCAGAAGGAC
TACAACGTTGCTGTTCGCATCACTCCCCTAAAATACGCCCAGATTTGCCTCTGGATTCACGG
GAACGATGCCAATTGTGCTTACGGC**TAA**CAGAGACCTGAAACAGGGCGGTGTATCATCTAAA
TCACAGAGAAAACCAGCTCTGCTTACCGTAGTGAGATCACTTCATAGGTTATGCCTGGACTT
GAACTCTGTCAATAGCATTTCAACATTTTTCAAAATCAGGAGATTTTCGTCCATTTAAAAAA
TGTATAGGTGCAGATATTGAAACTAGGTGGGCACTTCAATGCCAAGTATATACTCTTCTTTA
CATGGTGATGAGTTTCATTTGTAGAAAAATTTTGTTGCCTTCTTAAAAATTAGACACACTTT
AAACCTTCAAACAGGTATTATAAATAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCC
TACTCTAAGAAGAATCTAATAGGATGCTGGTTGTGTATTAAATGTGAAATTGCATAGATAAA
GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTACAACACAGTTTGTACTA
CTCTGAGATGGATCCATTCAGCTCATGCCCTCAATGTTTATATTGTGTTATCTGTTGGGTCT
GGGACATTTAGTTTAGTTTTTTTGAAGAATTACAAATCAGAAGAAAAGCAAGCATTATAAA
CAAAACTAATAACTGTTTTACTGCTTTAAGAAATAACAATTACAATGTGTATTATTTAAAAA
TGGGAGAAATAGTTTGTTCTATGAAATAAACCTAGTTTAGAAATAGGGAAGCTGAGACATTT
TAAGATCTCAAGTTTTTATTTAACTAATACTCAAAATATGGACTTTTCATGTATGCATAGGG
AAGACACTTCACAAATTATGAATGATCATGTGTTGAAAGCCACATTATTTTATGCTATACAT
TCTATGTATGAGGTGCTACATTTTTAGGACAAAGAATTCTGTAATCTTTTTCAAGAAAGAGT
CTTTTTCTCCTTGACAAAATCCAGCTTTTGTATGAGGACTATAGGGTGAATTCTCTGATTAG
TAATTTTAGATATGTCCTTTCCTAAAAATGAATAAAATTTATGAATATGA

FIGURE 118

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253

<subunit 1 of 1, 413 aa, 1 stop

<MW: 47070, pI: 9.92, NX(S/T): 3

MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFEADAKMMVNTVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNITTKGVSVRRKRQV
YGTDSRFSILDKRFLTNFPFSTAVKLSTGCSGILISPQHVLTAAHCVHDGKDYVKGSKKLRV
GLLKMRNKSGGKKRRGSKRSRREASGGDQREGTREHLQERAKGGRRKKSGRGQRIAEGRPS
FQWTRVKNTHIPKGWARGGMGDATLDYDYALLELKRAHKKKYMELGISPTIKKMPGGMIHFS
GFDNDRADQLVYRFCSVSDESNDLLYQYCDAESGSTGSGVYLRLKDPDKKNWKRKIIAVYSG
HQWVDVHGVQKDYNVAVRITPLKYAQICLWIHGNDANCAYG

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTTAGCACCAGTACTGGAT

GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTCAGTGTCCGATTCTGATTCCGGCAAGG

ATCCAAGC**ATG**GAATGCTGCCGTCGGGCAACTCCTGGCACACTGCTCCTCTTTCTGGCTTTC

CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTG

GGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCCTCCTACTCTCTGAGGCGCT

GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC

TGCCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA

CCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTCACTCA

AGTGCCAAGCCAAAGGAACAACCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACG

CGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGTTGGCTGCGA

TCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCA

CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAAATCGGATGATACT

GTGGTTGCACTTCCCTATGGAAGTAGACATATTCGCCTTGTCTTAAAAGGTCCTGATCACTT

ATATCTGGAAACCAAAACCCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAA

CTTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTTCCAGACAAAGAGATACTGAGA

ATGGCTGGACCACTCACAGCAGATTTCATTGTCAAGATTCGTAACTCGGGCTCCGCTGACAG

TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTC

CTTGCTCAGCAACCTGTGGAGGAGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG

AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACC

CAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGC

CTTATGACCTCTACCATCCCCTTCCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCC

TCGTGTGGGGGGGGCATCCAGAGCCGGGCAGTTTCCTGTGTGGAGGAGGACATCCAGGGGCA

TGTCACTTCAGTGGAAGAGTGGAAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCT

GCAACATTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGT

GGCCAGGGCCTCAGATACCGTGTGGTCCTCTGCATCGACCATCGAGGAATGCACACAGGAGG

CTGTAGCCCAAAAACAAAGCCCCACATAAAGAGGAATGCATCGTACCCACTCCCTGCTATA

AACCCAAAGAGAAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA

GAAGAAGGAGCTGCTGTGTCAGAGGAGCCCTCG**TAA**GTTGTAAAAGCACAGACTGTTCTATA

TTTGAAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGA

ACTAAGTGTAATCATCTCACCAAAGCTTTTTGGCTCTCAAATTAAAGATTGATTAGTTTCAA

AAAAAAAAA

FIGURE 120

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847

<subunit 1 of 1, 525 aa, 1 stop

<MW: 58416, pI: 6.62, NX(S/T): 1

MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRCLS
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQ
AKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFL
VDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCS
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKLPVEAKLPWFKQAQELEEGAAVSEEPS

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

FIGURE 121

CGGACGCGTGGGCGGCGGCTGCGGAACTCCCGTGGAGGGGCCGGTGGGCCCTCGGGCCTGAC

AG**ATG**GCAGTGGCCACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCG

GCCCGCCGGTTCGTGGGGCCCAGGGTCCAGCGGCTGCGCAGAGGCGGGGACCCCGGCCTCAT

GCACGGGAAGACTGTGCTGATCACCGGGGCGAACAGCGGCCTGGGCCGCGCCACGGCCGCCG

AGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCGCGCCGAGGAG

GCGGCGGGTCAGCTCCGCCGCGAGCTCCGCCAGGCCGCGGAGTGCGGCCCAGAGCCTGGCGT

CAGCGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCG

CCTTCTGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGG

ATCTTCCAGTGCCCTTACATGAAGACTGAAGATGGGTTTGAGATGCAGTTCGGAGTGAACCA

TCTGGGGCACTTTCTACTCACCAATCTTCTCCTTGGACTCCTCAAAAGTTCAGCTCCCAGCA

GGATTGTGGTAGTTTCTTCCAAACTTTATAAATACGGAGACATCAATTTTGATGACTTGAAC

AGTGAACAAAGCTATAATAAAGCTTTTGTTATAGCCGGAGCAAACTGGCTAACATTCTTTT

TACCAGGGAACTAGCCCGCCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTTGCATCCTG

GTATTGTACGGACAAATCTGGGGAGGCACATACACATTCCACTGTTGGTCAAACCACTCTTC

AATTTGGTGTCATGGGCTTTTTTCAAAAACTCCAGTAGAAGGTGCCCAGACTTCCATTTATTT

GGCCTCTTCACCTGAGGTAGAAGGAGTGTCAGGAAGATACTTTGGGGATTGTAAAGAGGAAG

AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAACTCTGGGATATCAGTGAAGTG

ATGGTTGGCCTGCTAAAA**TAG**GAACAAGGAGTAAAGAGCTGTTTATAAAACTGCATATCAG

TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACTTGAAGAAAAAGAATTTTG

ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTTGGAGTTACTGAAAAATTATTTT

TGGGATAAGAGAATTTCAGCAAGATGTTTTAAATATATATAGTAAGTATAATGAATAATAA

GTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCAAGCATGGATGACATATTA

ATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGTT

TCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTTTGTGTGGAAATTATCTGC

CTGGTGTGTGCACACAAGTCTTACTTGGAATAAATTTACTGGTAC

FIGURE 122

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747

<subunit 1 of 1, 336 aa, 1 stop

<MW: 36865, pI: 9.15, NX(S/T): 2

MAVATAAAVLAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEEAAGQLRRELRQAAECGPEPGVSGVGELIVRELDLASLRSVRA
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNHLGHFLLTNLLLGLLKSSAPSR
IVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSRSKLANILFTRELARRLEGTNVTVNVLHPG
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK

Important features:
Signal peptide:
amino acids 1-21

Short-chain alcohol dehydrogenase family protein
amino acids 134-144, 44-56 and 239-248

N-glycosylation site.
amino acids 212-215 and 239-242

FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN

TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTT

TATAAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACT

TGAAGAAAAGAATTTTGATATTGGAATAGCCTGNTAAGAGGNACATGTGGGTATTTTGGAG

TTACTGAAAAATTATTTTTGGGATAAGAGAATTTCAGCAAAGATGTTTTAAATATATATAGT

AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCA

AGCATGGATGACATATTAATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTT

TTCAAGTATCTTTGAGTTTCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTT

TGTGTGGAAATTATCTGCCTGGCTT

FIGURE 124

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCGGCTCCCGGAGCCCAGCC
CTTTCCTAACCCAACCCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTCACGGAC
CCCAGCGTTACC**ATG**CATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCT
GCTCCTGGTAACTTGGGTTTTTACTCCTGTAACAACTGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTTTGAGGAAGCTTCCGATGTCATTAAGGAAGAATT
TCCAAATGAAAATCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC
AGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAG
AGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGCAACAAAAAAG
TGACCCCATTCAAGAAATTCGGGACTTAGCAGAAATCACCACTCTTGATCGCAGCAAAAGAA
ATATCATTGGATATTTTGAGCAAAAGGACTCGGACAACTATAGAGTTTTTGAACGAGTAGCG
AATATTTTGCATGATGACTGTGCCTTTCTTTCTGCATTTGGGGATGTTTCAAAACCGGAAAG
ATATAGTGGCGACAACATAATCTACAAACCACCAGGGCATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTTTGATGTGACTTACAATTGGATTCAAGATAAATGTGTTCCTCTT
GTCCGAGAAATAACATTTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCAT
ACTCTTTCACATGAAAGAAGATACAGAAAGTTTAGAAATATTCCAGAATGAAGTAGCTCGGC
AATTAATAAGTGAAAAAGGTACAATAAACTTTTTACATGCCGATTGTGACAAATTTAGACAT
CCTCTTCTGCACATACAGAAAACTCCAGCAGATTGTCCTGTAATCGCTATTGACAGCTTTAG
GCATATGTATGTGTTTGGAGACTTCAAAGATGTATTAATTCCTGGAAAACTCAAGCAATTCG
TATTTGACTTACATTCTGGAAAACTGCACAGAGAATTCCATCATGGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTT**TAA**AAACTTG
AAAAACAGTTTGTAAGCCTTTCAACAGCAGCATCAACCTACGTGGTGGAAATAGTAAACCTA
TATTTTCATAATTCTATGTGTATTTTTATTTTGAATAAACAGAAAGAAATTTAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46927, pI: 5.21, NX(S/T): 0

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ

MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMKREYR

GQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKDSDNYRVFERVANILH

DDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREI

TFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH

IQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREFHHGPDPTDTAPG

EQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

FIGURE 126

ATTAAGGAAGAATTTCCAAATGAAAATCAAGTAGTNTTTGCCAGAGTNGATTGTGATCAGCA

CTCTGACATAGCCCAGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATG

GGATGATGATGAAGAGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTA

FIGURE 127

AGAGGCCTCTCTGGAAGTTGTCCCGGGTGTTCGCCGCNGGAGCCCGGGTCGAGAGGACNAGG

TGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCGGCTCCCGGAGCCCAGCCCTTTCCTAACCC

AACCCAACCTAGCCCNGTCCCAGCCGCCAGCGCCTGTCCCTGTCNCGGANCCCAGCGTNACC

ATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCTCCTGGTAAC

TTGGGTTTTTACTCCTGTAACAACTGAAATAACNNGTCTTGATACNNAGAATATAGATGAAA

TTTTAAACNATGCTGATGTGGCTTTAGTCAATTTTTATGCTGACTGGTGTCGTTTCAGTCAG

ATGTGGCATCCAATTTTTGAGGANGCTTCCGATGTCATTAAGGAAGAATTTCCAAATGAAAA

TCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA

TAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATACAGG

GGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGC

FIGURE 128

GCCCACGCGTCCG**ATG**GCGTTCACGTTCGCGGCCTTCTGCTACATGCTGGCGCTGCTGCTCA

CTGCCGCGCTCATCTTCTTCGCCATTTGGCACATTATAGCATTTGATGAGCTGAAGACTGAT

TACAAGAATCCTATAGACCAGTGTAATACCCTGAATCCCCTTGTACTCCCAGAGTACCTCAT

CCACGCTTTCTTCTGTGTCATGTTTCTTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA

TGCCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGATGAGTGGCCCAGGA

CTCTATGACCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG

GTGCAAATTAGCTTTTTATCTTCTAGCATTTTTTTACTACCTATATGGCATGATCTATGTTT

TGGTGAGCTCT**TAG**AACAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAAGCCAC

CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT

TACTTTAAAAAATGACTCCTTATTTTTTAAATGTTTCCACATTTTTGCTTGTGGAAAGACTG

TTTTCATATGTTATACTCAGATAAAGATTTTAAATGGTATTACGTATAAATTAATATAAAAT

GATTACCTCTGGTGTTGACAGGTTTGAACTTGCACTTCTTAAGGAACAGCCATAATCCTCTG

AATGATGCATTAATTACTGACTGTCCTAGTACATTGGAAGCTTTTGTTTATAGGAACTTGTA

GGGCTCATTTTGGTTTCATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC

TTCTGATGAAGTGAAAATGTATATCTGACTAGTGGGAAACTTCATGGGTTTCCTCATCTGTC

ATGTCGATGATTATATATGGATACATTTACAAAAATAAAAAGCGGGAATTTTCCCTTCGCTT

GAATATTATCCCTGTATATTGCATGAATGAGAGATTTCCCATATTTCCATCAGAGTAATAAA

TATACTTGCTTTAATTCTTAAGCATAAGTAAACATGATATAAAAATATATGCTGAATTACTT

GTGAAGAATGCATTTAAAGCTATTTTAAATGTGTTTTTATTTGTAAGACATTACTTATTAAG

AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCTTAAGAATTTGCAGG

TACTACAGATTTTCAAAACTGAATGAGAGAAAATTGTATAACCATCCTGCTGTTCCTTTAGT

GCAATACAATAAAACTCTGAAATTAAGACTC

FIGURE 129

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330

<subunit 1 of 1, 144 aa, 1 stop

<MW: 16699, pI: 5.60, NX(S/T): 0

MAFTFAAFCYMLALLLTAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHAFF
CVMFLCAAEWLTLGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS

Important features:

Signal peptide:

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

FIGURE 130

ATTATAGCATTTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAATACCCTG

AATCCCCTTGTACTCCCAGAGTACCTCATCCACGCTTTCTTCTGTGTCATGTTTCTTTGTGC

AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCCTCTTGGCATATCATATTTGGAGGTATA

TGAGTAGACCAGTGATGAGTGGCCCAGGACTCTATGACCCTACAACCATCATGAATGCAGAT

ATTCTAGCATATTGTCAGAAGGAAGGATGGTGCAAATTAGCTTTTTATCTTCTAGCATTTTT

TTACTACCTATATGGCATGATCTATGTTTTGGTGAGCTCTTAGAACAACACACAGAAGAATT

GGTCCAGTTAAGTGCATGCAAAAAGCCACCAAATGAAGGGATTCTATCCAGCAAGATCCTGT

CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTTAAAAAATG

FIGURE 131

CGGACGCGTGGGGGAAACCCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG

GAACAAG**ATG**GCGGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCGCCGC

TGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCGGGGACCGCTTCGGCTGAAGCATTTGAC

TCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACAC

CTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTC

AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA

GAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCC

ATTCGCTGAACTGAGACAAGAACAACTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTC

CTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC

TCTTCATGGACTTTTTATCTTCAAGCCGATGACGGAAAAATAGTTATATTCCAGTCTAAGCC

AGAAATCCAGTACGCACCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAA

GCAAAATGTCCTATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTTCTTGAAGATGGA

GAAAGTGATGGCTTTTTAAGATGCCTCTCTCTTAACTCTGGGTGGATTTTAACTACAACTCT

TGTCCTCTCGGTGATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGC

AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAG

CTAAACAGATATCCAGCTTCTTCTCTTGTGGTTGTTAGATCTAAAACTGAAGATCATGAAGA

AGCAGGGCCTCTACCTACAAAAGTGAATCTTGCTCATTCTGAAAT**TTA**AGCATTTTTCTTTT

AAAAGACAAGTGTAATAGACATCTAAAATTCCACTCCTCATAGAGCTTTTAAAATGGTTTCA

TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAATAAAGTTACTCAAATCTGTG

FIGURE 132

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847
<subunit 1 of 1, 323 aa, 1 stop
<MW: 36223, pI: 5.06, NX(S/T): 1
MAAPKGSLWVRTQLGLPPLLLLTMALAGGSGTASAEAFDSVLGDTASCHRACQLTYPLHTYP
KEEELYACQRGCRLFSICQFVDDGIDLNRTKLECESACTEAYSQSDEQYACHLGCQNQLPFA
ELRQEQLMSLMPKMHLLFPLTLVRSFWSDMMDSAQSFITSSWTFYLQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSYLQMRNSQAHRNFLEDGESDGFLRCLSLNSGWILTTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG
PLPTKVNLAHSEI
```

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

FIGURE 133

TTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACACCTACCC

TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTCAGTTTG

TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA

TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCATTCGC

TGAACTGAGACAAGAACAACTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCTAA

CTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGC

FIGURE 134

CACACTGGCCGGATCTTTTAGAGTCCTTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGGAACAAGATGGCGGCGCCGAAGGGAGCCTTTGGGTGAGGAC
CCAACTGGGGCTCCCGCCGCTGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCGGGGACCG
CTTCGGCTGAAGCATTTGACTCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAG
TTGACCTACCCCTTGCACACCTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTTCAATTTGTCAGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTT
GGTTGCCAGAATCAGCTGCCATTCGCTGAACTGAGACAAGAACAACTTATGTCCCTGATGCC
AAAAATGCACCTACTCTTTCCTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACT
CCGC

FIGURE 135

GCGAGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCCTGGGAGGCGGCCCGGAGGT

GGGGCGCCGCTGGGGCCGGCCCGCACGGGCTTCATCTGAGGGCGCACGGCCCGCGACCGAGC

GTGCGGACTGGCCTCCCAAGCGTGGGGCGACAAGCTGCCGGAGCTGCA**ATG**GGCCGCGGCTG

GGGATTCTTGTTTGGCCTCCTGGGCGCCGTGTGGCTGCTCAGCTCGGGCCACGGAGAGGAGC

AGCCCCCGGAGACAGCGGCACAGAGGTGCTTCTGCCAGGTTAGTGGTTACTTGGATGATTGT

ACCTGTGATGTTGAAACCATTGATAGATTTAATAACTACAGGCTTTTCCCAAGACTACAAAA

ACTTCTTGAAAGTGACTACTTTAGGTATTACAAGGTAAACCTGAAGAGGCCGTGTCCTTTCT

GGAATGACATCAGCCAGTGTGGAAGAAGGGACTGTGCTGTCAAACCATGTCAATCTGATGAA

GTTCCTGATGGAATTAAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA

AGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATCTCTGAGTGAGGAAACACAGA

AGGCTGTTCTTCAGTGGACCAAGCATGATGATTCTTCAGATAACTTCTGTGAAGCTGATGAC

ATTCAGTCCCCTGAAGCTGAATATGTAGATTTGCTTCTTAATCCTGAGCGCTACACTGGTTA

CAAGGGACCAGATGCTTGGAAAATATGGAATGTCATCTACGAAGAAAACTGTTTTAAGCCAC

AGACAATTAAAAGACCTTTAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGAACACT

TTTTACAGTTGGCTAGAAGGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGG

CCTACATGCAAGCATTAATGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAG

AAAAGAAATGGGGACACAACATTACAGAATTTCAACAGCGATTTGATGGAATTTTGACTGAA

GGAGAAGGTCCAAGAAGGCTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTT

ATCCAAAGTGTTACCATTCTTCGAGCGCCCAGATTTTCAACTCTTTACTGGAAATAAAATTC

AGGATGAGGAAAACAAAATGTTACTTCTGGAAATACTTCATGAAATCAAGTCATTTCCTTTG

CATTTTGATGAGAATTCATTTTTTGCTGGGGATAAAAAAGAAGCACACAAACTAAAGGAGGA

CTTTCGACTGCATTTTAGAAATATTTCAAGAATTATGGATTGTGTTGGTTGTTTTAAATGTC

GTCTGTGGGGAAAGCTTCAGACTCAGGGTTTGGGCACTGCTCTGAAGATCTTATTTTCTGAG

AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATTCCATCTAACCAGACAAGA

AATAGTATCATTATTCAACGCATTTGGAAGAATTTCTACAAGTGTGAAAGAATTAGAAAACT

TCAGGAACTTGTTACAGAATATTCAT**TAA**AGAAAACAAGCTGATATGTGCCTGTTTCTGGAC

AATGGAGGCGAAAGAGTGGAATTTCATTCAAAGGCATAATAGCAATGACAGTCTTAAGCCAA

ACATTTTATATAAAGTTGCTTTTGTAAAGGAGAATTATATTGTTTTAAGTAAACACATTTTT

AAAAATTGTGTTAAGTCTATGTATAATACTACTGTGAGTAAAAGTAATACTTTAATAATGTG

GTACAAATTTTAAAGTTTAATATTGAATAAAAGGAGGATTATCAAATTAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAA

FIGURE 136

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974

<subunit 1 of 1, 468 aa, 1 stop

<MW: 54393, pI: 5.63, NX(S/T): 2

MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAAQRCFCQVSGYLDDCTCDVETIDRFNNYRLF
PRLQKLLESDYFRYYKVNLKRPCPFWNDISQCGRRDCAVKPCQSDEVPDGIKSASYKYSEEA
NNLIEECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLLNPE
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLI
ELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFKCRLWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFH
LTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAAACTGTTTTAAGCCACAGACAATTAAAAGACCTT

TAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTTACAGTTGGCTAGAA

GGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA

TGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA

ACATTACAGAATTTNAACAGCGATTTGATGGAATTTTGACTGAAGGAGAAGGTCCAAGAAGG

CTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTTACCATT

CTTNGAGCGCCCAGATTTTCAACTNTTTACTGGAAATAAAATTCAGGATGAGGNAAACAAAA

TGTTACTTTTGGAAATACTTCATGAAATCAAGTCATTTCCTTTGCATTTTGATGAGAATTCA

TTTTTTTGCTG

FIGURE 138

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTTGGGAGGGGGCAGGATGGGAGGGAA

AGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAAC

CGATCAGGC**ATG**GAACTCCCCTTCGTCACTCACCTGTTCTTGCCCCTGGTGTTCCTGACAGG

TCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCTATTCCCAGGGCCACCAGAAG

CTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGATGCTGGTGGGC

GCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTTTATCGCTGCCCTGTAGGGGG

GGCCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTCATCTC

ATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGACAGATGGTGATGGGGGATTC

ATGGTGAGC**TAA**GGAGAGGGTGGTGGCAGTGTCTCTGAAGGTCCATAAAAGAAAAAAGAGAA

GTGTGGTAAGGGAAAATGGTCTGTGTGGAGGGGTCAAGGAGTTAAAAACCCTAGAAAGCAAA

AGGTAGGTAATGTCAGGGAGTAGTCTTCATGCCTCCTTCAACTGGGAGCATGTTCTGAGGGT

GCCCTCCCAAGCCTGGGAGTAACTATTTCCCCATCCCCAGGCCTGTGCCCCTCTCTGGTCT

CGTGCTTGTGGCAGCTCTGTCTTCAGTTCTGGGATATGTGCCCGTGTGGATGCTTCATTCCA

GCCTCAGGGAAGCCTGGCACCCACTGCCCAACGTGAGCCAGAGGAAGGCTGAGTACTTGGTT

CCCAGAAGGAGATACTGGGTGGGAAAAAGATGGGGCAAAGCGGTATGATGCCTGGCAAAGGG

CCTGCATGGCTATCCTCATTGCTACCTAATGTGCTTGCAAAAGCTCCATGTTTCCTAACAGA

TTCAGACTCCTGGCCAGGTGTGGTGGCCCACACCTGTAATTCTAGCACTTTGGGAGGCCAAG

GTGGGCAGATCACTTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACTCCAT

CTCTACTAAAAAAAAAAAAAATACAAAAATTAGCTGGGTGCGCTAGTGCATGCCTGTAATCTC

ATCTACTCGGGAGGCTAAGACAGGAGACTCTCACTTCAACCCAGGAGGTGGAGGTTGCGGTG

AGCCAAGATTGTGCCTCTGCACTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAAA

AATAATAATAATAATAATTCAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA

CTCATGCCTGTAATCCCAACATTTTGGGAGGCCAACGCAGGAGGATTGCTTGAGGTCTGGAG

GTTTGAGACCAGCCTGGGCAACATAGAAAGACCCCATCTCTAAATAAATGTTTTAAAAAT

FIGURE 139

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039

><subunit 1 of 1, 124 aa, 1 stop

><MW: 13352, pI: 5.99, NX(S/T): 1

MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPW

DGPSGDRRGDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETDGDGGFMVS

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

FIGURE 140

CACAGTTCCCCACCATCACTCNTCCCATTCCTTCCAACTTTATTTTTAGCTTGCCATTGGGA

GGGGGCAGGATGGGAGGGAAAGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGAC

TTCTCATACTGGACAGAAACCGATCAGGCATGGAACTCCCCTTCGTCACTCACCTGTTCTTG

CCCCTGGTGTTCCTGACAGGTCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCT

ATTCCCAGGGCCACCAGAAGCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGAC

AGCGATGGATGCTGGTGGGCGCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTT

TATCGCTGCCCTGTAGGGGGGGCCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTA

CCAACTGGGAAATTCATCTCATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGA

CAGATGGTGATGG

FIGURE 141

AAAGTTACATTTTCTCTGGAACTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACTGAGTCTACCA
A**ATG**CAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA
AACAGTGTACTATTCTGTCGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAAATGGTGAGGAGTGG
GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
CATTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT
GGTCGTGCCACTGTTCGTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG
TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCCAGAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT
CTCA**TAG**GTTTGCGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTTGCAAGGCTAGAGGGAAACTGGTGACACTCTACAGTCTGACTGATTCAG
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGGAGAGCCCACTTTCCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACTTCAAGCCCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACTACTGTAAAGTGTTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTTGTGCTCCTTTTTTC
TGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA
AAAAAAAAAA

FIGURE 142

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033

<subunit 1 of 1, 311 aa, 1 stop

<MW: 35076, pI: 5.04, NX(S/T): 2

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLLMWSPVIAPGE

TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTSAW

SILKHPFNRNSTILTRPGMEITKDGFHLVIELEDLGPQFEFLVAYWRREPGAEEHVKMVRSG

GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMLILV

VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

FIGURE 143

TCCTGCTGATGCACATCTGGGTTTGGCAAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT

CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA

TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG

TCAAACTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT

TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC

CTGCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA

GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTCGAATACCAGGGGGAGTACGAGAGCCT

GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG

ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGCCACATTGGGC

TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTAC

CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG

GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG

GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAGAGAGGCCCAATATGACCC

ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

FIGURE 144

CCCACGCGTCCGCCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA

GGCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGA

GGAGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAAGGAGAGAGGGAGGAGGAG

GAGGAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAAC

GGAGAGGAGGTGTGGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGA

GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGG

AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTG

GCTGCTTTGGCATTTGGGGAACTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCCT

GGAGGGACAGGGTCCCAGAAGGAGGGGACAGAGGAGCTGAGAGAGGGGGGCAGGGCGTTGGG

CAGGGGTCCCTCGGAGGCCTCCTGGGG**ATG**GGGGCTGCAGCTCGTCTGAGCGCCCCTCGAGC

GCTGGTACTCTGGGCTGCACTGGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCGAGG

ACTGGTGGAGCTACAAGGATAATCTCCAGGGAAACTTCGTGCCAGGGCCTCCTTTCTGGGGC

CTGGTGAATGCAGCGTGGAGTCTGTGTGCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGA

GCTGAAGAGGGTTCTTTATGACCCCTTTCTGCCCCCATTAAGGCTCAGCACTGGAGGAGAGA

AGCTCCGGGGAACCTTGTACAACACCGGCCGACATGTCTCCTTCCTGCCTGCACCCCGACCT

GTGGTCAATGTGTCTGGAGGTCCCCTCCTTTACAGCCACCGACTCAGTGAACTGCGGCTGCT

GTTTGGAGCTCGCGACGGAGCCGGCTCGGAACATCAGATCAACCACCAGGGCTTCTCTGCTG

AGGTGCAGCTCATTCACTTCAACCAGGAACTCTACGGGAATTTCAGCGCTGCCTCCCGCGGC

CCCAATGGCCTGGCCATTCTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAACCCATTCCT

CAGTCGCCTCCTTAACCGCGACACCATCACTCGCATCTCCTACAAGAATGATGCCTACTTTC

TTCAAGACCTGAGCCTGGAGCTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGC

TCTCTCAGCACCCCGCCCTGCTCCGAGACTGTCACCTGGATCCTCATTGACCGGGCCCTCAA

TATCACCTCCCTTCAGATGCACTCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCT

TCCAGAGCCTCAGCGGTAACAGCCGGCCCCTGCAGCCCTTGGCCCACAGGGCACTGAGGGGC

AACAGGGACCCCCGGCACCCCGAGAGGCGCTGCCGAGGCCCCAACTACCGCCTGCATGTGGA

TGGTGTCCCCCATGGTCGC**TGA**GACTCCCCTTCGAGGATTGCACCCGCCCGTCCTAAGCCTC

CCCACAAGGCGAGGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

FIGURE 145

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353

<subunit 1 of 1, 328 aa, 1 stop

<MW: 36238, pI: 9.90, NX(S/T): 3

MGAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC

AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPL

LYSHRLSELRLLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL

FVNVASTSNPFLSRLLNRDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSTPPCSE

TVTWILIDRALNITSLQMHSLRLLSQNPPSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER

RCRGPNYRLHVDGVPHGR

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

FIGURE 146

GGCGCCTGGTTCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGGTCGCCGCCAGCCTCCGCCGCCGAGCCTC
GTTCGTGTCCCCGCCCCTCGCTCCTGCAGCTACTGCTCAGAAACGCTGGGGCGCCCACCCTGGCAGACTAACGAA
GCAGCTCCCTTCCCACCCCAACTGCAGGTCTAATTTTGGACGCTTTGCCTGCCATTTCTTCCAGGTTGAGGGAGC
CGCAGAGGCGGAGGCTCGCGTATTCCTGCAGTCAGCACCCACGTCGCCCCCGGACGCTCGGTGCTCAGGCCCTTC
GCGAGCGGGGCTCTCCGTCTGCGGTCCCTTGTGAAGGCTCTGGGCGGCTGCAGAGGCCGGCCGTCCGGTTTGGCT
CACCTCTCCCAGGAAACTTCACACTGGAGAGCCAAAAGGAGTGGAAGAGCCTGTCTTGGAGATTTTCCTGGGGAA
ATCCTGAGGTCATTCATT**ATG**AAGTGTACCGCGCGGGAGTGGCTCAGAGTAACCACAGTGCTGTTCATGGCTAGA
GCAATTCCAGCCATGGTGGTTCCCAATGCCACTTTATTGGAGAAACTTTTGGAAAAATACATGGATGAGGATGGT
GAGTGGTGGATAGCCAAACAACGAGGGAAAAGGGCCATCACAGACAATGACATGCAGAGTATTTTGGACCTTCAT
AATAAATTACGAAGTCAGGTGTATCCAACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTGGAAAGA
TCTGCAGAATCCTGGGCTGAAAGTTGCTTGTGGGAACATGGACCTGCAAGCTTGCTTCCATCAATTGGACAGAAT
TTGGGAGCACACTGGGGAAGATATATAGGCCCCCGACGTTTCATGTACAATCGTGGTATGATGAAGTGAAAGACTTT
AGCTACCCATATGAACATGAATGCAACCCATATTGTCCATTCAGGTGTTCTGGCCCTGTATGTACACATTATACA
CAGGTCGTGTGGGCAACTAGTAACAGAATCGGTTGTGCCATTAATTTGTGTCATAACATGAACATCTGGGGGCAG
ATATGGCCCAAAGCTGTCTACCTGGTGTGCAATTACTCCCCAAAGGGAAACTGGTGGGGCCATGCCCCTTACAAA
CATGGGCGGCCCTGTTCTGCTTGCCCACCTAGTTTTGGAGGGGGCTGTAGAGAAAATCTGTGCTACAAAGAAGGG
TCAGACAGGTATTATCCCCCTCGAGAAGAGGAAACAAATGAAATAGAACGACAGCAGTCACAAGTCCATGACACC
CATGTCCGGACAAGATCAGATGATAGTAGCAGAAATGAAGTCATAAGCGCACAGCAAATGTCCCAAATTGTTTCT
TGTGAAGTAAGATTAAGAGATCAGTGCAAAGGAACAACCTGCAATAGGTACGAATGTCCTGCTGGCTGTTTGGAT
AGTAAAGCTAAAGTTATTGGCAGTGTACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCAATTCATTATGGT
ATAATAGACAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAAGCATTATTTCATCAAGTCCAATAGA
AATGGTATTCAAACAATTGGCAAATATCAGTCTGCTAATTCCTTCACAGTCTCTAAAGTAACAGTTCAGGCTGTG
ACTTGTGAAACAACTGTGGAACAGCTCTGTCCATTTCATAAGCCTGCTTCACATTGCCCAAGAGTATACTGTCCT
CGTAACTGTATGCAAGCAAATCCACATTATGCTCGTGTAATTGGAACTCGAGTTTATTCTGATCTGTCCAGTATC
TGCAGAGCAGCAGTACATGCTGGAGTGGTTCGAAATCACGGTGGTTATGTTGATGTAATGCCTGTGGACAAAAGA
AAGACCTACATTGCTTCTTTTCAGAATGGAATCTTCTCAGAAAGTTTACAGAATCCTCCAGGAGGAAAGGCATTC
AGAGTGTTTGCTGTTGTG**TGA**AACTGAATACTTGGAAGAGGACCATAAAGACTATTCCAAATGCAATATTTCTGA
ATTTTGTATAAAACTGTAACATTACTGTACAGAGTACATCAACTATTTTCAGCCCAAAAAGGTGCCAAATGCATA
TAAATCTTGATAAACAAAGTCTATAAAATAAAACATGGGACATTAGCTTTGGGAAAGTAATGAAAATATAATGG
TTTTAGAAATCCTGTGTTAAATATTGCTATATTTTTCTTAGCAGTTATTTCTACAGTTAATTACATAGTCATGATT
GTTCTACGTTTCATATATATTATATGGTGCTTTGTATATGCCACTAATAAAATGAATCTAAACATTGAATGTGAATG
GCCCTCAGAAAATCATCTAGTGCATTTAAAAATAATCGACTCTAAAACTGAAAGAAACCTTATCACATTTTCCCC
AGTTCAATGCTATGCCATTACCAACTCCAAATAATCTCAAATAATTTTCCACTTAATAACTGTAAAGTTTTTTTTC
TGTTAATTTAGGCATATAGAATATTAAAATTCTGATATTGCACTTCTTATTTTATATAAAATAATCCTTTAATATC
CAAATGAATCTGTTAAAATGTTTGATTCCTTGGGAATGGCCTTAAAAATAAATGTAATAAAGTCAGAGTGGTGGT
ATGAAACATTCCTAGTGATCATGTAGTAAATGTAGGGTTAAGCATGGACAGCCAGAGCTTTCTATGTACTGTTA
AAATTGAGGTCACATATTTTCTTTTGTATCCTGGCAAATACTCCTGCAGGCCAGGAAGTATAATAGCAAAAAGTT
GAACAAAGATGAACTAATGTATTACATTACCATTGCCACTGATTTTTTTTAAATGGTAAATGACCTTGTATATAA
ATATTGCCATATCATGGTACCTATAATGGTGATATATTTGTTTCATGAAAAATGTATTGTGCTTTGATACTAAA
AATCTGTAAAATGTTAGTTTTGGTAATTTTTTTTCTGCTGGTGGATTTACATATTAAATTTTTTCTGCTGGTGGA
TAAACATTAAAATTAATCATGTTTCAAAAAAAAAAAAA

FIGURE 147

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417

<subunit 1 of 1, 500 aa, 1 stop

<MW: 56888, pI: 8.53, NX(S/T): 2

MKCTAREWLRVTTVLFMARAIPAMVVPNATLLEKLLEKYMDEDGEWWIAKQRGKRAITDNDM

QSILDLHNKLRSQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHW

GRYRPPTFHVQSWYDEVKDFSYPYEHECNPYCPFRCSGPVCTHYTQVVWATSNRIGCAINLC

HNMNIWGQIWPKAVYLVCNYSPKGNWWGHAPYKHGRPCSACPPSFGGGCRENLCYKEGSDRY

YPPREEETNEIERQQSQVHDTHVRTRSDDSSRNEVISAQQMSQIVSCEVRLRDQCKGTTCNR

YECPAGCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDNDGGWVDITRQGRKHYFIKSNRNGI

QTIGKYQSANSFTVSKVTVQAVTCETTVEQLCPFHKPASHCPRVYCPRNCMQANPHYARVIG

TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSESLQNPPGGKAFRV

FAVV

Important features:

Signal peptide:

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

FIGURE 148

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGGAAGGCGCGCCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCCGGGCCGCCGCTCCGACGGGCCAGCGCCCTCCCC**ATG**TCCCTGCTCCCACGCCG
CGCCCCTCCGGTCAGCATGAGGCTCCTGGCGGCCGCGCTGCTCCTGCTGCTGCTGGCGCTGT
ACACCGCGCGTGTGGACGGGTCCAAATGCAAGTGCTCCCGGAAGGGACCCAAGATCCGCTAC
AGCGACGTGAAGAAGCTGGAAATGAAGCCAAAGTACCCGCACTGCGAGGAGAAGATGGTTAT
CATCACCACCAAGAGCGTGTCCAGGTACCGAGGTCAGGAGCACTGCCTGCACCCCAAGCTGC
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAA
GAA**TAG**GGTGAAAAACCTCAGAAGGGAAAACTCCAAACCAGTTGGGAGACTTGTGCAAAGGA
CTTTGCAGATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCCTTTC
TTTCTCACAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTTACCAACGGTCAG
TTTTTACATTTTATAGCTGCGTGCGAAAGGCTTCCAGATGGGAGACCCATCTCTCTTGTGCT
CCAGACTTCATCACAGGCTGCTTTTTATCAAAAAGGGGAAAACTCATGCCTTTCCTTTTTAA
AAAATGCTTTTTTGTATTTGTCCATACGTCACTATACATCTGAGCTTTATAAGCGCCCGGGA
GGAACAATGAGCTTGGTGGACACATTTCATTGCAGTGTTGCTCCATTCCTAGCTTGGGAAGC
TTCCGCTTAGAGGTCCTGGCGCCTCGGCACAGCTGCCACGGGCTCTCCTGGGCTTATGGCCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCCTGTAGCCGGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTCTGAGGAACTCAAGTTTGGTTGCCAGAAAAATGTGCTTCATTCCCCCCT
GGTTAATTTTTACACACCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGGTCTTTCCCAACCTGAGGATTTCTGAAAGGTTCACAGGTTCAATA
TTTAATGCTTCAGAAGCATGTGAGGTTCCCAACACTGTCAGCAAAAACCTTAGGAGAAAACT
TAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA
AACCTTCAAAGCATGTTTCTTTCCCTCACCACAACAGAACATGCAGTACTAAAGCAATATAT
TTGTGATTCCCCATGTAATTCTTCAATGTTAAACAGTGCAGTCCTCTTTCGAAAGCTAAGAT
GACCATGCGCCCTTTCCTCTGTACATATACCCTTAAGAACGCCCCCTCCACACACTGCCCCC
CAGTATATGCCGCATTGTACTGCTGTGTTATATGCTATGTACATGTCAGAAACCATTAGCAT
TGCATGCAGGTTTCATATTCTTTCTAAGATGGAAGTAATAAAATATATTTGAAATGTAAAA
AAAAAAAAAA

FIGURE 149

MSLLPRRAPPVSMRLLAAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH

CEEKMVIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNAWNEKRRVYEE

Signal sequence:

amino acids 1-34

FIGURE 150

GCCCCAGGGACTGCTATGGCTTCCTTTGTTGTTCACCCCGGTCTGCGTC**ATG**TTAAACTCCAATGTCCTCCTGTG
GTTAACTGCTCTTGCCATCAAGTTCACCCTCATTGACAGCCAAGCACAGTATCCAGTTGTCAACACAAATTATGG
CAAAATCCGGGGCCTAAGAACACCGTTACCCAATGAGATCTTGGGTCCAGTGGAGCAGTACTTAGGGGTCCCCTA
TGCCTCACCCCCCACTGGAGAGAGGCGGTTTCAGCCCCCAGAACCCCCGTCCTCCTGGACTGGCATCCGAAATAC
TACTCAGTTTGCTGCTGTGTGCCCCCAGCACCTGGATGAGAGATCCTTACTGCATGACATGCTGCCCATCTGGTT
TACCGCCAATTTGGATACTTTGATGACCTATGTTCAAGATCAAAATGAAGACTGCCTTTACTTAAACATCTACGT
GCCCACGGAAGATGGAGCCAACACAAAGAAAAACGCAGATGATATAACGAGTAATGACCGTGGTGAAGACGAAGA
TATTCATGATCAGAACAGTAAGAAGCCCGTCATGGTCTATATCCATGGGGGATCTTACATGGAGGGCACCGGCAA
CATGATTGACGGCAGCATTTTGGCAAGCTACGGAAACGTCATCGTGATCACCATTAACTACCGTCTGGGAATACT
AGGGTTTTTAAGTACCGGTGACCAGGCAGCAAAAGGCAACTATGGGCTCCTGGATCAGATTCAAGCACTGCGGTG
GATTGAGGAGAATGTGGGAGCCTTTGGCGGGGACCCCAAGAGAGTGACCATCTTTGGCTCGGGGGCTGGGGCCTC
CTGTGTCAGCCTGTTGACCCTGTCCCACTACTCAGAAGGTCTCTTCCAGAAGGCCATCATTCAGAGCGGCACCGC
CCTGTCCAGCTGGGCAGTGAACTACCAGCCGGCCAAGTACACTCGGATATTGGCAGACAAGGTCGGCTGCAACAT
GCTGGACACCACGGACATGGTAGAATGCCTGCGGAACAAGAACTACAAGGAGCTCATCCAGCAGACCATCACCCC
GGCCACCTACCACATAGCCTTCGGGCCGGTGATCGACGGCGACGTCATCCCAGACGACCCCCAGATCCTGATGGA
GCAAGGCGAGTTCCTCAACTACGACATCATGCTGGGCGTCAACCAAGGGGAAGGCCTGAAGTTCGTGGACGGCAT
CGTGGATAACGAGGACGGTGTGACGCCCAACGACTTTGACTTCTCCGTGTCCAACTTCGTGGACAACCTTTACGG
CTACCCTGAAGGGAAAGACACTTTGCGGGAGACTATCAAGTTCATGTACACAGACTGGGCCGATAAGGAAAACCC
GGAGACGCGGCGGAAAAACCCTGGTGGCTCTCTTTACTGACCACCAGTGGGTGGCCCCGCCGTGGCCGCCGACCT
GCACGCGCAGTACGGCTCCCCCACCTACTTCTATGCCTTCTATCATCACTGCCAAAGCGAAATGAAGCCCAGCTG
GGCAGATTCGGCCCATGGTGATGAGGTCCCCTATGTCTTCGGCATCCCCATGATCGGTCCCACCGAGCTCTTCAG
TTGTAACTTTTCCAAGAACGACGTCATGCTCAGCGCCGTGGTCATGACCTACTGGACGAACTTCGCCAAAACTGG
TGATCCAAATCAACCAGTTCCTCAGGATACCAAGTTCATTCACACAAAACCCAACCGCTTTGAAGAAGTGGCCTG
GTCCAAGTATAATCCCAAAGACCAGCTCTATCTGCATATTGGCTTGAAACCCAGAGTGAGAGATCACTACCGGGC
AACGAAAGTGGCTTTCTGGTTGGAACTCGTTCCTCATTTGCACAACTTGAACGAGATATTCCAGTATGTTTCAAC
AACCACAAAGGTTCCTCCACCAGACATGACATCATTTCCCTATGGCACCCGGCGATCTCCCGCCAAGATATGGCC
AACCACAAACGCCCAGCAATCACTCCTGCCAACAATCCCAAACACTCTAAGGACCCTCACAAAACAGGGCCTGA
GGACACAACTGTCCTCATTGAAACCAAACGAGATTATTCCACCGAATTAAGTGTCACCATTGCCGTCGGGGGCGTC
GCTCCTCTTCCTCAACATCTTAGCTTTTGCGGCGCTGTACTACAAAAAGGACAAGAGGCGCCATGAGACTCACAG
GCGCCCCAGTCCCCAGAGAAACACCACAAATGATATCGCTCACATCCAGAACGAAGAGATCATGTCTCTGCAGAT
GAAGCAGCTGGAACACGATCACGAGTGTGAGTCGCTGCAGGCACACGACACACTGAGGCTCACCTGCCCGCCAGA
CTACACCCTCACGCTGCGCCGGTCGCCAGATGACATCCCACTTATGACGCCAAACACCATCACCATGATTCCAAA
CACACTGACGGGGATGCAGCCTTTGCACACTTTTAACACCTTCAGTGGAGGACAAAACAGTACAAATTTACCCCA
CGGACATTCCACCACTAGAGTA**TAG**CTTTGCCCTATTTCCCTTCCTATCCCTCTGCCCTACCCGCTCAGCAACAT
AGAAGAGGGAAGGAAAGAGAGAAGGAAAGAGAGAGAGAAAGAAAGTCTCCAGACCAGGAATGTTTTTGTCCCACT
GACTTAAGACAAAAATGCAAAAAGGCAGTCATCCCATCCCGGCAGACCCTTATCGTTGGTGTTTTCCAGTATTAC
AAGATCAACTTCTGACCCTGTGAAATGTGAGAAGTACACATTTCTGTTAAAATAACTGCTTTAAGATCTCTACCA
CTCCAATCAATGTTTAGTGTGATAGGACATCACCATTTCAAGGCCCCGGGTGTTTCCAACGTCATGGAAGCAGCT
GACACTTCTGAAACTCAGCCAAGGACACTTGATATTTTTTAATTACAATGGAAGTTTAAACATTTCTTTCTGTGC
CACACAATGGATGGCTCTCCTTAAGTGAAGAAAGAGTCAATGAGATTTTGCCCAGCACATGGAGCTGTAATCCAG
AGAGAAGGAAACGTAGAAATTTATTATTAAAAGAATGGACTGTGCAGCGAAATCTGTACGGTTCTGTGCAAAGAG
GTGTTTTGCCAGCCTGAACTATATTTAAGAGACTTTGT

FIGURE 151

MLNSNVLLWLTALAIKFTLIDSQAQYPVVNTNYGKIRGLRTPLPNEILGPVEQYLGVPYASP

PTGERRFQPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMLPIWFTANLDTLMTYVQDQN

EDCLYLNIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMI

DGSILASYGNVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKR

VTIFGSGAGASCVSLLTLSHYSEGLFQKAIIQSGTALSSWAVNYQPAKYTRILADKVGCNML

DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPDDPQILMEQGEFLNYDIMLGV

NQGEGLKFVDGIVDNEDGVTPNDFDFSVSNFVDNLYGYPEGKDTLRETIKFMYTDWADKENP

ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMKPSWADSAHGDEVPYV

FGIPMIGPTELFSCNFSKNDVMLSAVVMTYWTNFAKTGDPNQPVPQDTKFIHTKPNRFEEVA

WSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHNLNEIFQYVSTTTKVPPPDMTS

FPYGTRRSPAKIWPTTKRPAITPANNPKHSKDPHKTGPEDTTVLIETKRDYSTELSVTIAVG

ASLLFLNILAFAALYYKKDKRRHETHRRPSPQRNTTNDIAHIQNEEIMSLQMKQLEHDHECE

SLQAHDTLRLTCPPDYTLTLRRSPDDIPLMTPNTITMIPNTLTGMQPLHTFNTFSGGQNSTN

LPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

FIGURE 152

GGGAAAG**ATG**GCGGCGACTCTGGGACCCCTTGGGTCGTGGCAGCAGTGGCGGCGATGTTTGT
CGGCTCGGGATGGGTCCAGGATGTTACTCCTTCTTCTTTTGTTGGGGTCTGGGCAGGGGCCA
CAGCAAGTCGGGGCGGGTCAAACGTTCGAGTACTTGAAACGGGAGCACTCGCTGTCGAAGCC
CTACCAGGGTGTGGGCACAGGCAGTTCCTCACTGTGGAATCTGATGGGCAATGCCATGGTGA
TGACCCAGTATATCCGCCTTACCCCAGATATGCAAAGTAAACAGGGTGCCTTGTGGAACCGG
GTGCCATGTTTCCTGAGAGACTGGGAGTTGCAGGTGCACTTCAAAATCCATGGACAAGGAAA
GAAGAATCTGCATGGGGATGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGC
CTGTGTTTGGAAACATGGACAAATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAAT
GAGGAGAAGCAGCAAGAGCGGGTATTCCCTACATCTCAGCCATGGTGAACAACGGCTCCCT
CAGCTATGATCATGAGCGGGATGGGCGGCCTACAGAGCTGGGAGGCTGCACAGCCATTGTCC
GCAATCTTCATTACGACACCTTCCTGGTGATTCGCTACGTCAAGAGGCATTTGACGATAATG
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCCGGAGTCCGCCTGCC
CCGCGGCTACTACTTCGGCACCTCCTCCATCACTGGGGATCTCTCAGATAATCATGATGTCA
TTTCCTTGAAGTTGTTTGAACTGACAGTGGAGAGAACCCCAGAAGAGGAAAAGCTCCATCGA
GATGTGTTCTTGCCCTCAGTGGACAATATGAAGCTGCCTGAGATGACAGCTCCACTGCCGCC
CCTGAGTGGCCTGGCCCTCTTCCTCATCGTCTTTTTCTCCCTGGTGTTTTCTGTATTTGCCA
TAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTAC**TGA**
GCCCTCCTGCTGCCACCACTTTTGTGACTGTCACCCATGAGGTATGGAAGGAGCAGGCACTG
GCCTGAGCATGCAGCCTGGAGAGTGTTCTTGTCTCTAGCAGCTGGTTGGGGACTATATTCTG
TCACTGGAGTTTTGAATGCAGGGACCCCGCATTCCCATGGTTGTGCATGGGGACATCTAACT
CTGGTCTGGGAAGCCACCCACCCCAGGGCAATGCTGCTGTGATGTGCCTTTCCCTGCAGTCC
TTCCATGTGGGAGCAGAGGTGTGAAGAGAATTTACGTGGTTGTGATGCCAAAATCACAGAAC
AGAATTTCATAGCCCAGGCTGCCGTGTTGTTTGACTCAGAAGGCCCTTCTACTTCAGTTTTG
AATCCACAAGAATTAAAAACTGGTAACACCACAGGCTTTCTGACCATCCATTCGTTGGGTT
TTGCATTTGACCCAACCCTCTGCCTACCTGAGGAGCTTTCTTTGGAAACCAGGATGGAAACT
TCTTCCCTGCCTTACCTTCCTTTCACTCCATTCATTGTCCTCTCTGTGTGCAACCTGAGCTG
GGAAAGGCATTTGGATGCCTCTCTGTTGGGGCCTGGGGCTGCAGAACACACCTGCGTTTCAC
TGGCCTTCATTAGGTGGCCCTAGGGAGATGGCTTTCTGCTTTGGATCACTGTTCCCTAGCAT
GGGTCTTGGGTCTATTGGCATGTCCATGGCCTTCCCAATCAAGTCTCTTCAGGCCCTCAGTG
AAGTTTGGCTAAAGGTTGGTGTAAAAATCAAGAGAAGCCTGGAAGACATCATGGATGCCATG
GATTAGCTGTGCAACTGACCAGCTCCAGGTTTGATCAAACCAAAAGCAACATTTGTCATGTG
GTCTGACCATGTGGAGATGTTTCTGGACTTGCTAGAGCCTGCTTAGCTGCATGTTTTGTAGT
TACGATTTTTGGAATCCCACTTTGAGTGCTGAAAGTGTAAGGAAGCTTTCTTCTTACACCTT
GGGCTTGGATATTGCCCAGAGAAGAAATTTGGCTTTTTTTTTCTTAATGGACAAGAGACAGT
TGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACCCTCATCATCTGTGCCTGGAAGAGTT
CACTGTCATTGAGCAGCACAGCCTGAGTGCTGGCCTCTGTCAACCCTTATTCCACTGCCTTA
TTTGACAAGGGGTTACATGCTGCTCACCTTACTGCCCTGGGATTAAATCAGTTACAGGCCAG
AGTCTCCTTGGAGGGCCTGGAACTCTGAGTCCTCCTATGAACCTCTGTAGCCTAAATGAAAT
TCTTAAAATCACCGATGGAACCAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCCATGG

FIGURE 153

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911

><subunit 1 of 1, 348 aa, 1 stop

><MW: 39711, pI: 8.70, NX(S/T): 1

MAATLGPLGSWQQWRRCLSARDGSRMLLLLLLLGSGQGPQQVGAGQTFEYLKREHSLSKPYQ

GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWELQVHFKIHGQGKKN

LHGDGLAIWYTKDRMQPGPVFGNMDKFVGLGVFVDTYPNEEKQQERVFPYISAMVNNGSLSY

DHERDGRPTELGGCTAIVRNLHYDTFLVIRYVKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG

YYFGTSSITGDLSDNHDVISLKLFELTVERTPEEEKLHRDVFLPSVDNMKLPEMTAPLPPLS

GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRKRFY

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

FIGURE 154

CCGAGCCGGGCGCGCAGCGACGGAGCTGGGGCCGGCCTGGGACCATGGGCGTGAGTGCAATCTACGGATCAGTCT
CTGATGGTGGGTCGTTAACCTCAGTGGGGACTCCAAGATTTCCATGAAGAAAATCAGTTGTCTTCATTCAAGAAT
TGGGGTCTGGCTCAGAATTCCTGCAGCTGGTGAAAATCTGTTTTCTAGAAGAGGTTTAATTAATGCCTGCAGTCT
GACATGTTCCCGATTTGAGGTGAAACCATGAAGAGAAAATAGAATACTTAATA**ATG**CTTTTCCGCAACCGCTTCT
TGCTGCTGCTGGCCCTGGCTGCGCTGCTGGCCTTTGTGAGCCTCAGCCTGCAGTTCTTCCACCTGATCCCGGTGT
CGACTCCTAAGAATGGAATGAGTAGCAAGAGTCGAAAGAGAATCATGCCCGACCCTGTGACGGGAGCCCCCTGTGA
CAGACCCCGTTTATGAAGCTCTTTTGTACTGCAACATCCCCAGTGTGGCCGAGCGCAGCATGGAAGGTCATGCCC
CGCATCATTTTAAGCTGGTCTCAGTGCATGTGTTCATTCGCCACGGAGACAGGTACCCACTGTATGTCATTCCCA
AAACAAAGCGACCAGAAATTGACTGCACTCTGGTGGCTAACAGGAAACCGTATCACCCAAAACTGGAAGCTTTCA
TTAGTCACATGTCAAAAGGATCCGGAGCCTCTTTCGAAAGCCCCTTGAACTCCTTGCCTCTTTACCCAAATCACC
CATTGTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCAGCATTTGCAGAACGGTCAGCTGCTGAGGGATA
TCTATCTAAAGAAACACAAACTCCTGCCCAATGATTGGTCTGCAGACCAGCTCTATTTAGAGACCACTGGGAAAA
GCCGGACCCTACAAAGTGGGCTGGCCTTGCTTTATGGCTTTCTCCCAGATTTTGACTGGAAGAAGATTTATTTCA
GGCACCAGCCAAGTGCGCTGTTCTGCTCTGGAAGCTGCTATTGCCCGGTAAGAAACCAGTATCTGGAAAAGGAGC
AGCGTCGTCAGTACCTCCTACGTTTGAAAAACAGCCAGCTGGAGAAGACCTACGGGGAGATGGCCAAGATCGTGG
ATGTCCCCACCAAGCAGCTTAGAGCTGCCAACCCCATAGACTCCATGCTCTGCCACTTCTGCCACAATGTCAGCT
TTCCCTGTACCAGAAATGGCTGTGTTGACATGGAGCACTTCAAGGTAATTAAGACCCATCAGATCGAGGATGAAA
GGGAAAGACGGGAGAAGAAATTGTACTTCGGGTATTCTCTCCTGGGTGCCCACCCCATCCTGAACCAAACCATCG
GCCGGATGCAGCGTGCCACCGAGGGCAGGAAAGAAGAGCTCTTTGCCCTCTACTCTGCTCATGATGTCACTCTGT
CACCAGTTCTCAGTGCCTTGGGCCTTTCAGAAGCCAGGTTCCCAAGGTTTGCAGCCAGGTTGATCTTTGAGCTTT
GGCAAGACAGAGAAAAGCCCAGTGAACATTCCGTCCGGATTCTTTACAATGGCGTCGATGTCACATTCCACACCT
CTTTCTGCCAAGACCACCACAAGCGTTCTCCCAAGCCCATGTGCCCGCTTGAAAACTTGGTCCGCTTTGTGAAAA
GGGACATGTTTGTAGCCCTGGGTGGCAGTGGTACAAATTATTATGATGCATGTCACAGGGAAGGATTC**TAA**AAGG
TATGCAGTACAGCAGTATAGAATCCATGCCAATACAGAGCATAGGGAAAGGTCCACTTCTAGTTTTGTCTGTTAC
TAAGGGTAGAAGATTATTGCTTTTTAAAGGCTAAATATTGTTTGTGGGAACCACAGATGGTTGGGGTTGAACAGT
AAGCACATTGCTGCAATGTGGTACGTGAATTGCTTGGTACAAAATGGCCAGTTCACAGAGGAATAGAAGGTACTT
TATCATAGCCAGACTTCGCTTAGAATGCCAGAATAATATAGTTCAAGACCTGAAGTTGCCAATCCAAGTTTGCAC
TCTTCTGGCCTGCCCCATGTTACTATGTGATGGAACCAGCACACCTCAACCAAAATTTTTTTAATCTTAGACATT
TTTACCTTGTCCTTGTTAAGAATTTCTTGAAGTGATTTATCTAAAATAAAGGTTGGCAAACTTTTTCTGTAAAGG
GCCAGATTGTAAATATTTCAGACTGTGTGGACCAAAAGGCCACATACAGTCTCTGTCATAACTACTCAACTCTGT
TTCTGAAGCAGGAAAGCCACCACAGACAGTACATAAAGGAATATGTGTAGCTGGGTTCCCAGGCCAGACAAAACA
GATGGTGACCAGACTTGGCCCCTGGGCTGTAGTTTGCTGACCCCTCATCTAAAAAATAGGCTATACTACAATTGC
ACTTCCAGCACTTTGAGAACGAGTTGAATACCAAGAATTATTCAATGGTTCCTCCAGTAACTTCTGCTAGAAACA
CAGAATTTGGTCTGTATCTGACACTAGAACAAAACTTGAGGGTAAATAAACATTGAATTAGAATGAATCATAGAA
AACTGATTAGAAGAATACTTGATGTTTATGATGATTGTGGTACAAGATAGTTTTAAGTATGTTCTAAATATTTGT
CTGCTGTAGTCTATTTGCTGTATATGCTGAAATTTTTGTATGCCATTTAGTATTTTTATAGTTTAGGAAAATATT
TTCTAAGACCAGTTTTAGATGACTCTTATTCCTGTAGTAATATTCAATTTGCTGTACCTGCTTGGTGGTTAGAAG
GAGGCTAGAAGATGAATTCAGGCACTTTCTTCCAATAAAACTAATTATGGCTCATTCCCTTTGACAAGCTGTAGA
ACTGGATTCATTTTTAAACCATTTTCATCAGTTTCAAATGGTAAATTCTGATTGATTTTTAAATGCGTTTTTGGA
AGAACTTTGCTATTAGGTAGTTTACAGATCTTTATAAGGTGTTTTATATATTAGAAGCAATTATAATTACATCTG
TGATTTCTGAACTAATGGTGCTAATTCAGAGAAATGGAAAGTGAAAGTGAGATTCTCTGTTGTCATCGGCATTCC
AACTTTTTCTCTTTGTTTTTGTCCAGTGTTGCATTTGAATATGTCTGTTTCTATAAATAAATTTTTTAAGAATAA

FIGURE 155

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329

><subunit 1 of 1, 480 aa, 1 stop

><MW: 55240, pI: 9.30, NX(S/T): 2

MLFRNRFLLLLALAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDPVTEPPVTDPVY

EALLYCNIPSVAERSMEGHAPHHFKLVSVHVFIRHGDRYPLYVIPKTKRPEIDCTLVANRKP

YHPKLEAFISHMSKGSGASFESPLNSLPLYPNHPLCEMGELTQTGVVQHLQNGQLLRDIYLK

KHKLLPNDWSADQLYLETTGKSRTLQSGLALLYGFLPDFDWKKIYFRHQPSALFCSGSCYCP

VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSMLCHFCHNVSFPC

TRNGCVDMEHFKVIKTHQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELF

ALYSAHDVTLSPVLSALGLSEARFPRFAARLIFELWQDREKPSEHSVRILYNGVDVTFHTSF

CQDHHKRSPKPMCPLENLVRFVKRDMFVALGGSGTNYYDACHREGF

Signal sequence:

amino acids 1-18

FIGURE 156

AAAAAAGCTCACTAAAGTTTCTATTAGAGCGAATACGGTAGATTTCCATCCCCTTTTGAAGAACAGTACTGTGGA
GCTATTTAAGAGATAAAAACGAAATATCCTTTCTGGGAGTTCAAGATTGTGCAGTAATTGGTTAGGACTCTGAGC
GCCGCTGTTCACCAATCGGGGAGAGAAAAGCGGAGATCCTGCTCGCCTTGCACGCGCCTGAAGCACAAAGCAGAT
AGCTAGGAATGAACCATCCCTGGGAGTATGTGGAAACAACGGAGGAGCTCTGACTTCCCAACTGTCCCATTCTAT
GGGCGAAGGAACTGCTCCTGACTTCAGTGGTTAAGGGCAGAATTGAAAATAATTCTGGAGGAAGATAAGA**ATG**AT
TCCTGCGCGACTGCACCGGGACTACAAAGGGCTTGTCCTGCTGGGAATCCTCCTGGGGACTCTGTGGGAGACCGG
ATGCACCCAGATACGCTATTCAGTTCCGGAAGAGCTGGAGAAAGGCTCTAGGGTGGGCGACATCTCCAGGGACCT
GGGGCTGGAGCCCCGGGAGCTCGCGGAGCGCGGAGTCCGCATCATCCCCAGAGGTAGGACGCAGCTTTTCGCCCT
GAATCCGCGCAGCGGCAGCTTGGTCACGGCGGGCAGGATAGACCGGGAGGAGCTCTGTATGGGGGCCATCAAGTG
TCAATTAAATCTAGACATTCTGATGGAGGATAAAGTGAAAATATATGGAGTAGAAGTAGAAGTAAGGGACATTAA
CGACAATGCGCCTTACTTTCGTGAAAGTGAATTAGAAATAAAAATTAGTGAAAATGCAGCCACTGAGATGCGGTT
CCCTCTACCCCACGCCTGGGATCCGGATATCGGGAAGAACTCTCTGCAGAGCTACGAGCTCAGCCCGAACACTCA
CTTCTCCCTCATCGTGCAAAATGGAGCCGACGGTAGTAAGTACCCCGAATTGGTGCTGAAACGCGCCCTGGACCG
CGAAGAAAAGGCTGCTCACCACCTGGTCCTTACGGCCTCCGACGGGGGCGACCCGGTGCGCACAGGCACCGCGCG
CATCCGCGTGATGGTTCTGGATGCGAACGACAACGCACCAGCGTTTGCTCAGCCCGAGTACCGCGCGAGCGTTCC
GGAGAATCTGGCCTTGGGCACGCAGCTGCTTGTAGTCAACGCTACCGACCCTGACGAAGGAGTCAATGCGGAAGT
GAGGTATTCCTTCCGGTATGTGGACGACAAGGCGGCCCAAGTTTTCAAACTAGATTGTAATTCAGGGACAATATC
AACAATAGGGGAGTTGGACCACGAGGAGTCAGGATTCTACCAGATGGAAGTGCAAGCAATGGATAATGCAGGATA
TTCTGCGCGAGCCAAAGTCCTGATCACTGTTCTGGACGTGAACGACAATGCCCCAGAAGTGGTCCTCACCTCTCT
CGCCAGCTCGGTTCCCGAAAACTCTCCCAGAGGGACATTAATTGCCCTTTTAAATGTAAATGACCAAGATTCTGA
GGAAAACGGACAGGTGATCTGTTTCATCCAAGGAAATCTGCCCTTTAAATTAGAAAAATCTTACGGAAATTACTA
TAGTTTAGTCACAGACATAGTCTTGGATAGGGAACAGGTTCCTAGCTACAACATCACAGTGACCGCCACTGACCG
GGGAACCCCGCCCCTATCCACGGAAACTCATATCTCGCTGAACGTGGCAGACACCAACGACAACCCGCCGGTCTT
CCCTCAGGCCTCCTATTCCGCTTATATCCCAGAGAACAATCCCAGAGGAGTTTCCCTCGTCTCTGTGACCGCCCA
CGACCCCGACTGTGAAGAGAACGCCCAGATCACTTATTCCCTGGCTGAGAACACCATCCAAGGGGCAAGCCTATC
GTCCTACGTGTCCATCAACTCCGACACTGGGGTACTGTATGCGCTGAGCTCCTTCGACTACGAGCAGTTCCGAGA
CTTGCAAGTGAAAGTGATGGCGCGGGACAACGGGCACCCGCCCCTCAGCAGCAACGTGTCGTTGAGCCTGTTCGT
GCTGGACCAGAACGACAATGCGCCCGAGATCCTGTACCCCGCCCTCCCCACGGACGGTTCCACTGGCGTGGAGCT
GGCTCCCCGCTCCGCAGAGCCCGGCTACCTGGTGACCAAGGTGGTGGCGGTGGACAGAGACTCCGGCCAGAACGC
CTGGCTGTCCTACCGTCTGCTCAAGGCCAGCGAGCCGGGACTCTTCTCGGTGGGTCTGCACACGGGCGAGGTGCG
CACGGCGCGAGCCCTGCTGGACAGAGACGCGCTCAAGCAGAGCCTCGTAGTGGCCGTCCAGGACCACGGCCAGCC
CCCTCTCTCCGCCACTGTCACGCTCACCGTGGCCGTGGCCGACAGCATCCCCCAAGTCCTGGCGGACCTCGGCAG
CCTCGAGTCTCCAGCTAACTCTGAAACCTCAGACCTCACTCTGTACCTGGTGGTAGCGGTGGCCGCGGTCTCCTG
CGTCTTCCTGGCCTTCGTCATCTTGCTGCTGGCGCTCAGGCTGCGGCGCTGGCACAAGTCACGCCTGCTGCAGGC
TTCAGGAGGCGGCTTGACAGGAGCGCCGGCGTCGCACTTTGTGGGCGTGGACGGGGTGCAGGCTTTCCTGCAGAC
CTATTCCCACGAGGTTTCCCTCACCACGGACTCGCGGAAGAGTCACCTGATCTTCCCCCAGCCCAACTATGCAGA
CATGCTCGTCAGCCAGGAGAGCTTTGAAAAAAGCGAGCCCCTTTTGCTGTCAGGTGATTCGGTATTTTCTAAAGA
CAGTCATGGGTTAATTGAGGTGAGTTTATATCAAATCTTCTTTCTTTTTTTTTTTAATTGCTCTGTCTCCCAAGC
TGGAGTGCAGCGGTACGATCATAGCTCACTGCGGCCTCAAACTCCTAGGCTCAAGCAATTATCCCACCTTTGCCT
CCGGTGTAACAGGGACTACAGGTGCAAGCCACCTACTGTCTGCCTATCTATCTATCTATCTATCTATCTAT
CTATCTATCTATCTATCTATTACTTTCTTGTACAGACGGGAGTCTCACGCCTGTAATCCCAGTACTTTGGGAGGC
CGAGGCGGGTGGATCACCTGAGGTTGGGAGTTTGAGACCAGCC**TGA**CCAACATGGAGAAACCCCGTCTATACTAA
AAAAATACAAAATTAGCCGGGCGTGGTGGTGCATGTCTGTAATCCCAGCTACTTGGGAGGCTGAGTCAGGAGAAT
TGCTTTAACCTGGGAGGTGGAGGTTGCAATGAGCTGAGATTGTGCCATTGCACTCCAGCCTGGGCAACAAGAGTG
AAACTCTATCTCA

FIGURE 157

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306

><subunit 1 of 1, 916 aa, 1 stop

><MW: 100204, pI: 4.92, NX(S/T): 4

MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVPEELEKGSRVGDISRDLGLEPRELAER

GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVR

DINDNAPYFRESELEIKISENAATEMRFPLPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA

DGSKYPELVLKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFAQPEYR

ASVPENLALGTQLLVVNATDPDEGVNAEVRYSFRYVDDKAAQVFKLDCNSGTISTIGELDHE

ESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLNVN

DQDSEENGQVICFIQGNLPFKLEKSYGNYYSLVTDIVLDREQVPSYNITVTATDRGTPPLST

ETHISLNVADTNDNPPVFPQASYSAYIPENNPRGVSLVSVTAHDPDCEENAQITYSLAENTI

QGASLSSYVSINSDTGVLYALSSFDYEQFRDLQVKVMARDNGHPPLSSNVSLSLFVLDQNDN

APEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVDRDSGQNAWLSYRLLKASEPGLFSVG

LHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTLTVAVADSIPQVLADLGSLESPA

NSETSDLTLYLVVAVAAVSCVFLAFVILLLALRLRRWHKSRLLQASGGGLTGAPASHFVGVD

GVQAFLQTYSHEVSLTTDSRKSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFSKDSHGL

IEVSLYQIFFLFFFNCSVSQAGVQRYDHSSLRPQTPRLKQLSHLCLRCNRDYRCKPPTVCLS

IYLSIYLSIYLSIYLLLSCTDGSLTPVIPVLWEAEAGGSPEVGSLRPA

Signal sequence:
amino acids 1-30

Transmembrane domains:
amino acids 693-711, 809-823, 869-888

FIGURE 158

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAG

GCTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA

TCAGTAGGTGACCCCGCCCCTGGATTCTGGAAGACCTCACC**ATG**GGACGCCCCCGACCTCGT

GCGGCCAAGACGTGGATGTTCCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGC

ACAGGAGGACAAGGTGCTGGGGGGTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGG

CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAGGTGGCAACTGGGTCCTT

ACAGCTGCCCACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAA

TAAAGATGGCCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACA

GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCC

CTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTG

CACCGTCTCAGGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACT

GTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACA

GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGG

CCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGA

GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC

ATAGGCAGCAAGGGC**TGA**TTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACTCT

CTGGTTC

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336

<subunit 1 of 1, 260 aa, 1 stop

<MW: 28048, pI: 7.87, NX(S/T): 1

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVL

VGGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLL

QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCED

AYPGQITDGMVCAGSSKGADTCQGDSGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRY

LDWIKKIIGSKG

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

FIGURE 160

```
GGCGCCGGTGCACCGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCCTGCGCGCCCCGGCCCGC
CGCGCCGCCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCCGGGCCCGCGCCCGC
GCCCGCGCCCAGGTGAGCGCTCCGCCCGCCGCGAGGCCCCGCCCCGGCCCGCCCCCGCCCCG
CCCCGGCCGGCGGGGGAACCGGGCGGATTCCTCGCGCGTCAAACCACCTGATCCCATAAAAC
ATTCATCCTCCCGGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCCGCCGCCCTCG
CCCTGTGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCCAGCCAGAGCCGGGCGGAGC
GGAGCGCGCCGAGCCTCGTCCCGCGGCCGGGCCGGGGCCGGGCCGTAGCGGCGGCGCCTGGA
TGCGGACCCGGCCGCGGGGAGACGGGCGCCCGCCCCGAAACGACTTTCAGTCCCCGACGCGC
CCCGCCCAACCCCTACGATGAAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTG
CTGTGGCTGCAGGCCTGGCAGGTGGCAGCCCCATGCCCAGGTGCCTGCGTATGCTACAATGA
GCCCAAGGTGACGACAAGCTGCCCCCAGCAGGGCCTGCAGGCTGTGCCCGTGGGCATCCCTG
CTGCCAGCCAGCGCATCTTCCTGCACGGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTC
CGTGCCTGCCGCAACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCGAATTGATGC
GGCTGCCTTCACTGGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGGACCCTGCCACATTCCACGGCCTGGGCCGCCTACACACGCTGCACCTGGACCGC
TGCGGCCTGCAGGAGCTGGGCCCGGGGCTGTTCCGCGGCCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACAACGCGCTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCA
CACACCTCTTCCTGCACGGCAACCGCATCTCCAGCGTGCCCGAGCGCGCCTTCCGTGGGCTG
CACAGCCTCGACCGTCTCCTACTGCACCAGAACCGCGTGGCCCATGTGCACCCGCATGCCTT
CCGTGACCTTGGCCGCCTCATGACACTCTATCTGTTTGCCAACAATCTATCAGCGCTGCCCA
CTGAGGCCCTGGCCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTG
TGTGACTGCCGGGCACGCCCACTCTGGGCCTGGCTGCAGAAGTTCCGCGGCTCCTCCTCCGA
GGTGCCCTGCAGCCTCCCGCAACGCCTGGCTGGCCGTGACCTCAAACGCCTAGCTGCCAATG
ACCTGCAGGGCTGCGCTGTGGCCACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACC
GATGAGGAGCCGCTGGGGCTTCCCAAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGT
ACTGGAGCCTGGAAGACCAGCTTCGGCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCGGTG
ACAGCCCGCCGGGCAACGGCTCTGGCCCACGGCACATCAATGACTCACCCTTTGGGACTCTG
CCTGGCTCTGCTGAGCCCCCGCTCACTGCAGTGCGGCCCGAGGGCTCCGAGCCACCAGGGTT
CCCCACCTCGGGCCCTCGCCGGAGGCCAGGCTGTTCACGCAAGAACCGCACCCGCAGCCACT
GCCGTCTGGGCCAGGCAGGCAGCGGGGGTGGCGGGACTGGTGACTCAGAAGGCTCAGGTGCC
CTACCCAGCCTCACCTGCAGCCTCACCCCCCTGGGCCTGGCGCTGGTGCTGTGGACAGTGCT
TGGGCCCTGCTGACCCCCAGCGGACACAAGAGCGTGCTCAGCAGCCAGGTGTGTGTACATAC
GGGGTCTCTCTCCACGCCGCCAAGCCAGCCGGGCGGCCGACCCGTGGGGCAGGCCAGGCCAG
GTCCTCCCTGATGGACGCCTGCCGCCCGCCACCCCCATCTCCACCCCATCATGTTTACAGGG
TTCGGCGGCAGCGTTTGTTCCAGAACGCCGCCTCCCACCCAGATCGCGGTATATAGAGATAT
GCATTTTATTTTACTTGTGTAAAAATATCGGACGACGTGGAATAAAGAGCTCTTTTCTTAAA
AAAA
```

FIGURE 161

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184

><subunit 1 of 1, 473 aa, 1 stop

><MW: 50708, pI: 9.28, NX(S/T): 6

MKRASAGGSRLLAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRI

FLHGNRISHVPAASFRACRNLTILWLHSNVLARIDAAAFTGLALLEQLDLSDNAQLRSVDPA

TFHGLGRLHTLHLDRCGLQELGPGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH

GNRISSVPERAFRGLHSLDRLLLHQNRVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAP

LRALQYLRLNDNPWVCDCRARPLWAWLQKFRGSSSEVPCSLPQRLAGRDLKRLAANDLQGCA

VATGPYHPIWTGRATDEEPLGLPKCCQPDAADKASVLEPGRPASAGNALKGRVPPGDSPPGN

GSGPRHINDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA

GSGGGGTGDSEGSGALPSLTCSLTPLGLALVLWTVLGPC

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

FIGURE 162

GGAAGTCCACGGGGAGCTTGGATGCCAAAGGGAGGACGGCTGGGTCCTCTGGAGAGGACTAC
TCACTGGCATATTTCTGAGGTATCTGTAGAATAACCACAGCCTCAGATACTGGGGACTTTAC
AGTCCCACAGAACCGTCCTCCCAGGAAGCTGAATCCAGCAAGAACA**ATG**GAGGCCAGCGGGA
AGCTCATTTGCAGACAAAGGCAAGTCCTTTTTTCCTTTCTCCTTTTGGGCTTATCTCTGGCG
GGCGCGGCGGAACCTAGAAGCTATTCTGTGGTGGAGGAAACTGAGGGCAGCTCCTTTGTCAC
CAATTTAGCAAAGGACCTGGGTCTGGAGCAGAGGGAATTCTCCAGGCGGGGGGTTAGGGTTG
TTTCCAGAGGGAACAAACTACATTTGCAGCTCAATCAGGAGACCGCGGATTTGTTGCTAAAT
GAGAAATTGGACCGTGAGGATCTGTGCGGTCACACAGAGCCCTGTGTGCTACGTTTCCAAGT
GTTGCTAGAGAGTCCCTTCGAGTTTTTTCAAGCTGAGCTGCAAGTAATAGACATAAACGACC
ACTCTCCAGTATTTCTGGACAAACAAATGTTGGTGAAAGTATCAGAGAGCAGTCCTCCTGGG
ACTACGTTTCCTCTGAAGAATGCCGAAGACTTAGATGTAGGCCAAAACAATATTGAGAACTA
TATAATCAGCCCCAACTCCTATTTTCGGGTCCTCACCCGCAAACGCAGTGATGGCAGGAAAT
ACCCAGAGCTGGTGCTGGACAAAGCGCTGGACCGAGAGGAAGAAGCTGAGCTCAGGTTAACA
CTCACAGCACTGGATGGTGGCTCTCCGCCCAGATCTGGCACTGCTCAGGTCTACATCGAAGT
CCTGGATGTCAACGATAATGCCCCTGAATTTGAGCAGCCTTTCTATAGAGTGCAGATCTCTG
AGGACAGTCCGGTAGGCTTCCTGGTTGTGAAGGTCTCTGCCACGGATGTAGACACAGGAGTC
AACGGAGAGATTTCCTATTCACTTTTCCAAGCTTCAGAAGAGATTGGCAAAACCTTTAAGAT
CAATCCCTTGACAGGAGAAATTGAACTAAAAAAACAACTCGATTTCGAAAAACTTCAGTCCT
ATGAAGTCAATATTGAGGCAAGAGATGCTGGAACCTTTTCTGGAAAATGCACCGTTCTGATT
CAAGTGATAGATGTGAACGACCATGCCCCAGAAGTTACCATGTCTGCATTTACCAGCCCAAT
ACCTGAGAACGCGCCTGAAACTGTGGTTGCACTTTTCAGTGTTTCAGATCTTGATTCAGGAG
AAAATGGGAAAATTAGTTGCTCCATTCAGGAGGATCTACCCTTCCTCCTGAAATCCGCGGAA
AACTTTTACACCCTACTAACGGAGAGACCACTAGACAGAGAAAGCAGAGCGGAATACAACAT
CACTATCACTGTCACTGACTTGGGGACCCCTATGCTGATAACACAGCTCAATATGACCGTGC
TGATCGCCGATGTCAATGACAACGCTCCCGCCTTCACCCAAACCTCCTACACCCTGTTCGTC
CGCGAGAACAACAGCCCCGCCCTGCACATCCGCAGCGTCAGCGCTACAGACAGAGACTCAGG
CACCAACGCCCAGGTCACCTACTCGCTGCTGCCGCCCCAGGACCCGCACCTGCCCCTCACAT
CCCTGGTCTCCATCAACGCGGACAACGGCCACCTGTTCGCCCTCAGGTCTCTGGACTACGAG
GCCCTGCAGGGGTTCCAGTTCCGCGTGGGCGCTTCAGACCACGGCTCCCCGGCGCTGAGCAG
CGAGGCGCTGGTGCGCGTGGTGGTGCTGGACGCCAACGACAACTCGCCCTTCGTGCTGTACC
CGCTGCAGAACGGCTCCGCGCCCTGCACCGAGCTGGTGCCCCGGGCGGCCGAGCCGGGCTAC
CTGGTGACCAAGGTGGTGGCGGTGGACGGCGACTCGGGCCAGAACGCCTGGCTGTCGTACCA
GCTGCTCAAGGCCACGGAGCTCGGTCTGTTCGGCGTGTGGGCGCACAATGGCGAGGTGCGCA
CCGCCAGGCTGCTGAGCGAGCGCGACGCGGCCAAGCACAGGCTGGTGGTGCTGGTCAAGGAC
AATGGCGAGCCTCCGCGCTCGGCCACCGCCACGCTGCACGTGCTCCTGGTGGACGGCTTCTC
CCAGCCCTACCTGCCTCTCCCGGAGGCGGCCCCGACCCAGGCCCAGGCCGACTTGCTCACCG
TCTACCTGGTGGTGGCGTTGGCCTCGGTGTCTTCGCTCTTCCTCTTTTCGGTGCTCCTGTTC
GTGGCGGTGCGGCTGTGTAGGAGGAGCAGGGCGGCCTCGGTGGGTCGCTGCTTGGTGCCCGA
GGGCCCCCTTCCAGGGCATCTTGTGGACATGAGCGGCACCAGGACCCTATCCCAGAGCTACC
AGTATGAGGTGTGTCTGGCAGGAGGCTCAGGGACCAATGAGTTCAAGTTCCTGAAGCCGATT
ATCCCCAACTTCCCTCCCCAGTGCCCTGGGAAAGAAATACAAGGAAATTCTACCTTCCCCAA
TAACTTTGGGTTCAATATTCAG**TGA**CCATAGTTGACTTTTACATTCCATAGGTATTTTATTT
TGTGGCATTTCCATGCCAATGTTTATTTCCCCCAATTTGTGTGTATGTAATATTGTACGGAT
TTACTCTTGATTTTTCTCATGTTCTTTCTCCCTTTGTTTTAAAGTGAACATTTACCTTTATT
CCTGGTTCTT

FIGURE 163

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314

<subunit 1 of 1, 798 aa, 1 stop

<MW: 87552, pI: 4.84, NX(S/T): 5

MEASGKLICRQRQVLFSFLLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSRGNKLHLQLNQETADLLLNEKLDREDLCGHTEPCVLRFQVLLESPFEFFQAELQV
IDINDHSPVFLDKQMLVKVSESSPPGTTFPLKNAEDLDVGQNNIENYIISPNSYFRVLTRKR
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSGTAQVYIEVLDVNDNAPEFEQPFY
RVQISEDSPVGFLVVKVSATDVDTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSPIPENAPETVVALFSVS
DLDSGENGKISCSIQEDLPFLLKSAENFYTLLTERPLDRESRAEYNITITVTDLGTPMLITQ
LNMTVLIADVNDNAPAFTQTSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGFQFRVGASDHGSPALSSEALVRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELGLFGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAQ
ADLLTVYLVVALASVSSLFLFSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMSGTRT
LSQSYQYEVCLAGGSGTNEFKFLKPIIPNFPPQCPGKEIQGNSTFPNNFGFNIQ

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

FIGURE 164

ACCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCGCGTAGCCGTGC

GCCGATTGCCTCTCGGCCTGGGCA**ATG**GTCCCGGCTGCCGGTCGACGACCGCCCCGCGTCAT

GCGGCTCCTCGGCTGGTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCCGTCCGCGGCGTGG

AGGTTGCAGAGGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCCTCTCCAGGTG

GGGGCTGTGTACCTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGGCCAGGACAGGGCAGC

AGAAGAGGCCAATGCGGTGCTGGGGCTGGACACCCAAGGCGATCACATGGTGATGCTGTCTG

TGATTCCTGGGGAAGCTGAGGACAAAGTGAGTTCAGAGCCTAGCGGCGTCACCTGTGGTGCT

GGAGGAGCGGAGGACTCAAGGTGCAACGTCCGAGAGAGCCTTTTCTCTCTGGATGGCGCTGG

AGCACACTTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACG

CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACTGTGAG

GAGAGAAACATTACAGGATTAGAAAATTTCACTCTGAAAATTTTAAATATGTCACAGGACCT

TATGGATTTTCTGAACCCAAACGGTAGTGACTGTACTCTAGTCCTGTTTTACACCCCGTGGT

GCCGCTTTTCTGCCAGTTTGGCCCCTCACTTTAACTCTCTGCCCCGGGCATTTCCAGCTCTT

CACTTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTTCTACCAGGTTTGGCACCGTAGC

TGTTCCTAATATTTTATTATTTCAAGGAGCTAAACCAATGGCCAGATTTAATCATACAGATC

GAACACTGGAAACACTGAAAATCTTCATTTTTAATCAGACAGGTATAGAAGCCAAGAAGAAT

GTGGTGGTAACTCAAGCCGACCAAATAGGCCCTCTTCCCAGCACTTTGATAAAAGTGTGGA

CTGGTTGCTTGTATTTTCCTTATTCTTTTTAATTAGTTTTATTATGTATGCTACCATTCGAA

CTGAGAGTATTCGGTGGCTAATTCCAGGACAAGAGCAGGAACATGTGGAG**TAG**TGATGGTCT

GAAAGAAGTTGGAAAGAGGAACTTCAATCCTTCGTTTCAGAAATTAGTGCTACAGTTTCATA

CATTTTCTCCAGTGACGTGTTGACTTGAAACTTCAGGCAGATTAAAAGAATCATTTGTTGAA

CAACTGAATGTATAAAAAATTATAAACTGGTGTTTAACTAGTATTGCAATAAGCAAATGC

AAAAATATTCAATAG

FIGURE 165

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333

><subunit 1 of 1, 360 aa, 1 stop

><MW: 39885, pI: 4.79, NX(S/T): 7

MVPAAGRRPPRVMRLLGWWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGEE

ELLHDPMGQDRAAEEANAVLGLDTQGDHMVMLSVIPGEAEDKVSSEPSGVTCGAGGAEDSRC

NVRESLFSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCEERNITGLE

NFTLKILNMSQDLMDFLNPNGSDCTLVLFYTPWCRFSASLAPHFNSLPRAFPALHFLALDAS

QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTLETLKIFIFNQTGIEAKKNVVVTQADQ

IGPLPSTLIKSVDWLLVFSLFFLISFIMYATIRTESIRWLIPGQEQEHVE

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to dilsufide isomerase

amino acids 212-302

N-glycosylation site.

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296

Thioredoxin domain

amino acids 211-227

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTCGCGCGCCCACG**ATG**CTGCAGGGCCCTGGCT

CGCTGCTGCTGCTCTTCCTCGCCTCGCACTGCTGCCTGGGCTCGGCGCGCGGGCTCTTCCTC

TTTGGCCAGCCCGACTTCTCCTACAAGCGCAGCAATTGCAAGCCCATCCCGGTCAACCTGCA

GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCAACCTGCTGGGCCACGAGACCA

TGAAGGAGGTGCTGGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCG

GACACCAAGAAGTTCCTGTGCTCGCTCTTCGCCCCCGTCTGCCTCGATGACCTAGACGAGAC

CATCCAGCCATGCCACTCGCTCTGCGTGCAGGTGAAGGACCGCTGCGCCCCGGTCATGTCCG

CCTTCGGCTTCCCCTGGCCCGACATGCTTGAGTGCGACCGTTTCCCCCAGGACAACGACCTT

TGCATCCCCCTCGCTAGCAGCGACCACCTCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATG

TGAAGCCTGCAAAAATAAAAATGATGATGACAACGACATAATGGAAACGCTTTGTAAAAATG

ATTTTGCACTGAAAATAAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATC

CTGGAGACCAAGAGCAAGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAA

ATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGCACCTGTGAGGAGATGAACGACATCAACG

CGCCCTATCTGGTCATGGGACAGAAACAGGGTGGGGAGCTGGTGATCACCTCGGTGAAGCGG

TGGCAGAAGGGGCAGAGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGC**TA**

GTCCCGGCATCCTGATGGCTCCGACAGGCCTGCTCCAGAGCACGGCTGACCATTTCTGCTCC

GGGATCTCAGCTCCCGTTCCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGGCAGCT

TCCCCCTGCCTTTTGCACGTTTGCATCCCCAGCATTTCCTGAGTTATAAGGCCACAGGAGTG

GATAGCTGTTTTCACCTAAAGGAAAAGCCCACCCGAATCTTGTAGAAATATTCAAACTAATA

AAATCATGAATATTTTAA

FIGURE 167

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920

><subunit 1 of 1, 295 aa, 1 stop

><MW: 33518, pI: 7.74, NX(S/T): 0

MLQGPGSLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPVNLQLCHGIEYQNMRLPN

LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDR

CAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDDDNDIM

ETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLNGVSERDLKKSVLWLKDSLQCTCE

EMNDINAPYLVMGQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homolgous to frizzled N terminus

amino acids 6-153

FIGURE 168

GTGGAGGCCGCCGACG**ATG**GCGGGGCCGACGGAGGCCGAGACGGGGTTGGCCGAGCCCCGGG

CCCTGTGCGCGCAGCGGGGCCACCGCACCTACGCGCGCCGCTGGGTGTTCCTGCTCGCGATC

AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTTGCACCTGTGGCTGACGTCAT

TGCTGAGGACTTGGTCCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG

TATCCACCCCATTTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTCGGGCTCCGTGCGGCG

ACCATCCTGGGTGCGTGGCTGAACTTTGCCGGGAGTGTGCTACGCATGGTGCCCTGCATGGT

TGTTGGGACCCAAAACCCATTTGCCTTCCTCATGGGTGGCCAGAGCCTCTGTGCCCTTGCCC

AGAGCCTGGTCATCTTCTCTCCAGCCAAGCTGGCTGCCTTGTGGTTCCCAGAGCACCAGCGA

GCCACGGCCAACATGCTCGCCACCATGTCGAACCCTCTGGGCGTCCTTGTGGCCAATGTGCT

GTCCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCCGTTAATGCTCGGTGTCTATACCATCC

CTGCTGGCGTCGTCTGCCTGCTGTCCACCATCTGCCTGTGGGAGAGTGTGCCCCCCACCCCG

CCCTCTGCCGGGGCTGCCAGCTCCACCTCAGAGAAGTTCCTGGATGGGCTCAAGCTGCAGCT

CATGTGGAACAAGGCCTATGTCATCCTGGCTGTGTGCTTGGGGGGAATGATCGGGATCTCTG

CCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGGTTTTCC

GGCCTCTGTGGCGCTCTCTTCATCACGTTTGGGATCCTGGGGGCACTGGCTCTCGGCCCCTA

TGTGGACCGGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTCTGG

CCTGCGTGCCCTTTGCCCTGGTGTCCCAGCTGCAGGGACAGACCCTTGCCCTGGCTGCCACC

TGCTCGCTGCTCGGGCTGTTTGGCTTCTCGGTGGGCCCCGTGGCCATGGAGTTGGCGGTCGA

GTGTTCCTTCCCCGTGGGGGAGGGGGCTGCCACAGGCATGATCTTTGTGCTGGGGCAGGCCG

AGGGAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGGAGCCGTCCTTG

TCCACCTGCCAGCAGGGGGAGGATCCACTTGACTGGACAGTGTCTCTGCTGCTGATGGCCGG

CCTGTGCACCTTCTTCAGCTGCATCCTGGCGGTCTTCTTCCACACCCCATACCGGCGCCTGC

AGGCCGAGTCTGGGGAGCCCCCCTCCACCCGTAACGCCGTGGGCGGCGCAGACTCAGGGCCG

GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGGTCCTGGGGCCCAGCACGGCGACTCCGGA

GTGCACGGCGAGGGGGGCCTCGCTAGAGGACCCCAGAGGGCCCGGGAGCCCCCACCCAGCCT

GCCACCGAGCGACTCCCCGTGCGCAAGGCCCAGCAGCCACCGACGCGCCCTCCCGCCCCGGC

AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTTATTGACCCGGCTGGGTCTCACTCCTCCTT

CTCCTCCCCGTGGGTGATCACG**TAG**CTGAGCGCCTTGTAGTCCAGGTTGCCCGCCACATCGA

TGGAGGCGAACTGGAACATCTGGTCCACCTGCGGGCGGGGGCGAAAGGGCTCCTTGCGGGCT

CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988

><subunit 1 of 1, 560 aa, 1 stop

><MW: 58427, pI: 6.86, NX(S/T): 2

MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLV

LSMEQINWLSLVYLVVSTPFGVAAIWILDSVGLRAATILGAWLNFAGSVLRMVPCMVVGTQN

PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRATANMLATMSNPLGVLVANVLSPVLV

KKGEDIPLMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNKA

YVILAVCLGGMIGISASFSALLEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK

HFTEATKIGLCLFSLACVPFALVSQLQGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV

GEGAATGMIFVLGQAEGILIMLAMTALTVRRSEPSLSTCQQGEDPLDWTVSLLLMAGLCTFF

SCILAVFFHTPYRRLQAESGEPPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG

ASLEDPRGPGSPHPACHRATPRAQGPAATDAPSRPGRLAGRVQASRFIDPAGSHSSFSSPWVIT

Important features:
Signal peptide:
amino acids 1-44

Transmembrane domains:
amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268, 280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.
amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.
amino acids 468-471

FIGURE 170

GTCCCACATCCTGCTCAACTGGGTCAGGTCCCTCTTAGACCAGCTCTTGTCCATCATTTGCTGAAGTGGACCAAC
TAGTTCCCCAGTAGGGGGTCTCCCCTGGCAATTCTTGATCGGCGTTTGGACATCTCAGATCGCTTCCAATGAAGA
TGGCCTTGCCTTGGGGTCCTGCTTGTTTCATAATCATCTAACTATGGGACAAGGTTGTGCCGGCAGCTCTGGGGG
AAGGAGCACGGGGCTGATCAAGCCATCCAGGAAACACTGGAGGACTTGTCCAGCCTTGAAAGAACTCTAGTGGTT
TCTGAATCTAGCCCACTTGGCGGTAAGC**ATG**ATGCAACTTCTGCAACTTCTGCTGGGGCTTTTGGGGCCAGGTGG
CTACTTATTTCTTTTAGGGGATTGTCAGGAGGTGACCACTCTCACGGTGAAATACCAAGTGTCAGAGGAAGTGCC
ATCTGGTACAGTGATCGGGAAGCTGTCCCAGGAACTGGGCCGGGAGGAGAGGCGGAGGCAAGCTGGGGCCGCCTT
CCAGGTGTTGCAGCTGCCTCAGGCGCTCCCCATTCAGGTGGACTCTGAGGAAGGCTTGCTCAGCACAGGCAGGCG
GCTGGATCGAGAGCAGCTGTGCCGACAGTGGGATCCCTGCCTGGTTTCCTTTGATGTGCTTGCCACAGGGGATTT
GGCTCTGATCCATGTGGAGATCCAAGTGCTGGACATCAATGACCACCAGCCACGGTTTCCCAAAGGCGAGCAGGA
GCTGGAAATCTCTGAGAGCGCCTCTCTGCGAACCCGGATCCCCCTGGACAGAGCTCTTGACCCAGACACAGGCCC
TAACACCCTGCACACCTACACTCTGTCTCCCAGTGAGCACTTTGCCTTGGATGTCATTGTGGGCCCTGATGAGAC
CAAACATGCAGAACTCATAGTGGTGAAGGAGCTGGACAGGGAAATCCATTCATTTTTTGATCTGGTGTTAACTGC
CTATGACAATGGGAACCCCCCCAAGTCAGGTACCAGCTTGGTCAAGGTCAACGTCTTGGACTCCAATGACAATAG
CCCTGCGTTTGCTGAGAGTTCACTGGCACTGGAAATCCAAGAAGATGCTGCACCTGGTACGCTTCTCATAAAACT
GACCGCCACAGACCCTGACCAAGGCCCCAATGGGGAGGTGGAGTTCTTCCTCAGTAAGCACATGCCTCCAGAGGT
GCTGGACACCTTCAGTATTGATGCCAAGACAGGCCAGGTCATTCTGCGTCGACCTCTAGACTATGAAAAGAACCC
TGCCTACGAGGTGGATGTTCAGGCAAGGGACCTGGGTCCCAATCCTATCCCAGCCCATTGCAAAGTTCTCATCAA
GGTTCTGGATGTCAATGACAACATCCCAAGCATCCACGTCACATGGGCCTCCCAGCCATCACTGGTGTCAGAAGC
TCTTCCCAAGGACAGTTTTATTGCTCTTGTCATGGCAGATGACTTGGATTCAGGACACAATGGTTTGGTCCACTG
CTGGCTGAGCCAAGAGCTGGGCCACTTCAGGCTGAAAAGAACTAATGGCAACACATACATGTTGCTAACCAATGC
CACACTGGACAGAGAGCAGTGGCCCAAATATACCCTCACTCTGTTAGCCCAAGACCAAGGACTCCAGCCCTTATC
AGCCAAGAAACAGCTCAGCATTCAGATCAGTGACATCAACGACAATGCACCTGTGTTTGAGAAAAGCAGGTATGA
AGTCTCCACGCGGGAAAACAACTTACCCTCTCTTCACCTCATTACCATCAAGGCTCATGATGCAGACTTGGGCAT
TAATGGAAAAGTCTCATACCGCATCCAGGACTCCCCAGTTGCTCACTTAGTAGCTATTGACTCCAACACAGGAGA
GGTCACTGCTCAGAGGTCACTGAACTATGAAGAGATGCCGGCTTTGAGTTCCAGGTGATCGCAGAGGACAGCGG
GCAACCCATGCTTGCATCCAGTGTCTCTGTGTGGGTCAGCCTCTTGGATGCCAATGATAATGCCCCAGAGGTGGT
CCAGCCTGTGCTCAGCGATGGAAAAGCCAGCCTCTTCCGTGCTTGTGAATGCCTCCACAGGCCACCTGCTGGTGCC
CATCGAGACTCCCAATGGCTTGGGCCCAGCGGGCACTGACACACCTCCACTGGCCACTCACAGCTCCCGGCCATT
CCTTTTGACAACCATTGTGGCAAGAGATGCAGACTCGGGGGCAAATGGAGAGCCCCTCTACAGCATCCGCAATGG
AAATGAAGCCCACCTCTTCATCCTCAACCCTCATACGGGGCAGCTGTTCGTCAATGTCACCAATGCCAGCAGCCT
CATTGGGAGTGAGTGGGAGCTGGAGATAGTAGTAGAGGACCAGGGAAGCCCCCCCTTACAGACCCGAGCCCTGTT
GAGGGTCATGTTTGTCACCAGTGTGGACCACCTGAGGGACTCAGCCCGCAAGCCTGGGGCCTTGAGCATGTCGAT
GCTGACGGTGATCTGCCTGGCTGTACTGTTGGGCATCTTCGGGTTGATCCTGGCTTTGTTCATGTCCATCTGCCG
GACAGAAAAGAAGGACAACAGGGCCTACAACTGTCGGGAGGCCGAGTCCACCTACCGCCAGCAGCCCAAGAGGCC
CCAGAAACACATTCAGAAGGCAGACATCCACCTCGTGCCTGTGCTCAGGGGTCAGGCAGGTGAGCCTTGTGAAGT
CGGGCAGTCCCACAAAGATGTGGACAAGGAGGCGATGATGGAAGCAGGCTGGGACCCCTGCCTGCAGGCCCCCTT
CCACCTCACCCCGACCCTGTACAGGACGCTGCGTAATCAAGGCAACCAGGGAGCACCGGCGGAGAGCCGAGAGGT
GCTGCAAGACACGGTCAACCTCCTTTTCAACCATCCCAGGCAGAGGAATGCCTCCCGGGAGAACCTGAACCTTCC
CGAGCCCCAGCCTGCCACAGGCCAGCCACGTTCCAGGCCTCTGAAGGTTGCAGGCAGCCCCACAGGGAGGCTGGC
TGGAGACCAGGGCAGTGAGGAAGCCCCACAGAGGCCACCAGCCTCCTCTGCAACCCTGAGACGGCAGCGACATCT
CAATGGCAAAGTGTCCCCTGAGAAAGAATCAGGGCCCCGTCAGATCCTGCGGAGCCTGGTCCGGCTGTCTGTGGC
TGCCTTCGCCGAGCGGAACCCCGTGGAGGAGCTCACTGTGGATTCTCCTCCTGTTCAGCAAATCTCCCAGCTGCT
GTCCTTGCTGCATCAGGGCCAATTCCAGCCCAAACCAAACCACCGAGGAAATAAGTACTTGGCCAAGCCAGGAGG
CAGCAGGAGTGCAATCCCAGACACAGATGGCCCAAGTGCAAGGGCTGGAGGCCAGACAGACCCAGAACAGGAGGA
AGGGCCTTTGGATCCTGAAGAGGACCTCTCTGTGAAGCAACTGCTAGAAGAAGAGCTGTCAAGTCTGCTGGACCC
CAGCACAGGTCTGGCCCTGGACCGGCTGAGCGCCCCTGACCCGGCCTGGATGGCGAGACTCTCTTTGCCCCTCAC
CACCAACTACCGTGACAATGTGATCTCCCCGGATGCTGCAGCCACGGAGGAGCCGAGGACCTTCCAGACGTTCGG
CAAGGCAGAGGCACCAGAGCTGAGCCCAACAGGCACGAGGCTGGCCAGCACCTTTGTCTCGGAGATGAGCTCACT
GCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCCGTGGAGGCCGCCTCCGAGGCGCTGCGGCGGCTCTCGGT
CTGCGGGAGGACCCTCAGTTTAGACTTGGCCACCAGTGCAGCCTCAGGCATGAAAGTGCAAGGGACCCAGGTGG
AAAGACGGGGACTGAGGGCAAGAGCAGAGGCAGCAGCAGCAGCAGCAGGTGCCT**GTGA**ACATACCTCAGACGCCT
CTGGATCCAAGAACCAGGGGCCTGAGGATCTGTGGACAAGAGCTGGTTTCTAAAATCTTGTAACTCACTAGCTAG
CGGCGGCCTGAGAACTTTAGGGTGACTGATGCTACCCCCACAGAGGAGGCAAGAGCCCCAGGACTAACAGCTGAC
TGACCAAAGCAGCCCCTTGTAAGCAGCTCTGAGTCTTTTGGAGGACAGGGACGGTTTGTGGCTGAGATAAGTGTT
TCCTGGCAAAACATATGTGGAGCACAAAGGGTCAGTCCTCTGGCAGAACAGATGCCACGGAGTATCACAGGCAGG
AAAGGGTGGCCTTCTTGGGTAGCAGGAGTCAGGGGGCTGTACCCTGGGGGTGCCAGGAAATGCTCTCTGACCTAT
CAATAAAGGAAAAGCAGTAAAAAAAAAAAAAAAAAAAA

FIGURE 171

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331

<subunit 1 of 1, 1184 aa, 1 stop

<MW: 129022, pI: 5.20, NX(S/T): 5

MMQLLQLLLGLLGPGGYLFLLGDCQEVTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQA

GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQ

VLDINDHQPRFPKGEQELEISESASLRTRIPLDRALDPDTGPNTLHTYTLSPSEHFALDVIV

GPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSNDNSPAFAESS

LALEIQEDAAPGTLLIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPL

DYEKNPAYEVDVQARDLGPNPIPAHCKVLIKVLDVNDNIPSIHVTWASQPSLVSEALPKDSF

IALVMADDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQD

QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLGINGKVS

YRIQDSPVAHLVAIDSNTGEVTAQRSLNYEEMAGFEFQVIAEDSGQPMLASSVSVWVSLLDA

NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPPLATHSSRPFLLTT

IVARDADSGANGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGS

PPLQTRALLRVMFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEK

KDNRAYNCREAESTYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA

GWDPCLQAPFHLTPTLYRTLRNQGNQGAPAESREVLQDTVNLLFNHPRQRNASRENLNLPEP

QPATGQPRSRPLKVAGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQ

ILRSLVRLSVAAFAERNPVEELTVDSPPVQQISQLLSLLHQGQFQPKPNHRGNKYLAKPGGS

RSAIPDTDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPD

PAWMARLSLPLTTNYRDNVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSL

LEMLLEQRSSMPVEAASEALRRLSVCGRTLSLDLATSAASGMKVQGDPGGKTGTEGKSRGSS

SSSRCL

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 amd 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

FIGURE 172

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCAGTCCCGGCTGCAGCACCTGGGAGAAGG

CAGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAG

GCAGGAGCCTTCCTTACACTTCGCC**ATG**AGTTTCCTCATCGACTCCAGCATCATGATTACCT

CCCAGATACTATTTTTTGGATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTAT

GAGATACGTCAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGCACCAT

GTTTGAGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACT

GGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTCATGGTGCCTTTTTACATTGGC

TATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTTCCTGTCTCTT

ATGGCTGACCTTTATGTATTTCTTCTGGAAACTAGGAGATCCCTTTCCCATTCTCAGCCCAA

AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTGGTGTGATTGGAGTGACTCTC

ATGGCTCTTCTTTCTGGATTTGGTGCTGTCAACTGCCCATACACTTACATGTCTTACTTCCT

CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATA

TGATCATAAGCAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAA

GTGCATAACAAACCATCAGGTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGG

AAGTGAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGC

TTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTC

AAGGGGAAATATTTTAATTTTCTTGGTTACTTTTTCTCTATTTACTGTGTTTGGAAAATTTT

CATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGTCACAAGAGGCA

TTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTTGGTCCCAACACATT

TCCTTCATTCTTGTTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC

CAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCCTGCTATTAGCAC

AGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAA

TACCGCACCATAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTT

TGATGTGATCTTCCTGGTCAGCGCTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAAC

AGGCACCAGAGAAGCAAATGGCACCT**TGA**ACTTAAGCCTACTACAGACTGTTAGAGGCCAGT

GGTTTCAAAATTTAGATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAAC

AAACAAATGCTATGGTAGCATTTTTCACCTTCATAGCATACTCCTTCCCCGTCAGGTGATA

CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGAGAACTAACTCAAGACAATACTCA

GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGAGCCAAGAAACTAA

AGGTGAAAAATACACTGGAACTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGT

AGGATTTCCGTTTTAAGGTTCACATGGAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATT

AAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG

ACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEI

LGVLNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFF

WKLGDPFPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDI

LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQ

QEVDALEELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF

DRVGKTDPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISS

SKSSNVIVLLLAQIMGMYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSA

LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTTCCTTACACTCGCCATGAGTTTCCTCATCGACTCCAGCA

TCATGATTACCTCCCNGANACTATTTTTTGGATTTGGGTGGCTTTTCTTCNGCGCCAATGTT

TAAAGACTATGAGATACGTCAGTATGTTGTACNGGTGATCTTCTCCGTGACGTTTGCCATTT

CTTGCACCATGTTTGAGCTCATCATCTTTGAAATCTTNGGAGTATTGAATAGCAGCTCCCGT

TATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTNTCATGGTGCCTTT

TTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTT

CCTGTCTCTTATGGCTGACCTTTATGTATTTCCAG

FIGURE 175

GTGTTGCCCTTGGGGAGGGGAAGGGGAGCCNGGCCCTTTCCTAAAATTTGGCCAAGGGTTTC
TTTNTTGAATTCCGGGTTNNGNATACCTTCCCAGAAAATATTTTTTGGATTTGGGGTAGNTT
TTTTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAGGTGATNTT
NTCCGTGACGTTTGCATTTTCTTGCACCATGTTTGAGCTCATCATNTTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATC
CTGGTTTTCATGGTGCCTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCA
TAAACAACGACTGCTTTTTTCCTGTCTNTTATGGCTGACCTTTATGTATTTNTTNTGGAAAN
TAGGAGATCCCTTTCCCATTCTC

FIGURE 176

CTCGCGCAGGGATCGTCCC**ATG**GCCGGGGCTCGGAGCCGCGACCCTTGGGGGGCCTCCGGGATTTGCTACCTTTT
TGGCTCCCTGCTCGTCGAACTGCTCTTCTCACGGGCTGTCGCCTTCAATCTGGACGTGATGGGTGCCTTGCGCAA
GGAGGGCGAGCCAGGCAGCCTCTTCGGCTTCTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCCCAGAGCTG
GCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCCTGGGCAGCAGGCGAATCGCACTGGAGGCCTCTTCGCTTG
CCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAA
GGAGAACCAGTGGTTGGGAGTCAGTGTTCGGAGCCAGGGGCCTGGGGGCAAGATTGTTACCTGTGCACACCGATA
TGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTCGCTGCTTTGTGCTCAGCCAGGA
CCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGAAGTTCTGTGAGGGACGCCCCCAAGGCCATGAACAATT
TGGGTTCTGCCAGCAGGGCACAGCTGCCGCCTTCTCCCCTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAAC
CTATAATTGGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGGACGACGG
TCCCTACGAGGCGGGGGGAGAGAAGGAGCAGGACCCCCGCCTCATCCCGGTCCCTGCCAACAGCTACTTTGGCTT
CTCTATTGACTCGGGGAAAGGTCTGGTGCGTGCAGAAGAGCTGAGCTTTGTGGCTGGAGCCCCCCGCGCCAACCA
CAAGGGTGCTGTGGTCATCCTGCGCAAGGACAGCGCCAGTCGCCTGGTGCCCGAGGTTATGCTGTCTGGGGAGCG
CCTGACCTCCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGATGGCTGGCCAGACCTGATAGTGGG
TGCCCCCTACTTCTTTGAGCGCCAAGAAGAGCTGGGGGGTGCTGTGTATGTGTACTTGAACCAGGGGGGTCACTG
GGCTGGGATCTCCCCTCTCCGGCTCTGCGGCTCCCCTGACTCCATGTTCGGGATCAGCCTGGCTGTCCTGGGGGA
CCTCAACCAAGATGGCTTTCCAGATATTGCAGTGGGTGCCCCCTTTGATGGTGATGGGAAAGTCTTCATCTACCA
TGGGAGCAGCCTGGGGGTTGTCGCCAAACCTTCACAGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGG
CTACTCCCTGTCAGGCAGCTTGGATATGGATGGGAACCAATACCCTGACCTGCTGGTGGGCTCCCTGGCTGACAC
CGCAGTGCTCTTCAGGGCCAGACCCATCCTCCATGTCTCCCATGAGGTCTCTATTGCTCCACGAAGCATCGACCT
GGAGCAGCCCAACTGTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAGGGTCTGTTTCAGCTACATTGCAGTCCC
CAGCAGCTATAGCCCTACTGTGGCCCTGGACTATGTGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGT
TCCCCGTGTGACGTTCCTGAGCCGTAACCTGGAAGAACCCAAGCACCAGGCCTCGGGCACCGTGTGGCTGAAGCA
CCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTCGGGCCATTGT
AGTGACCTTGTCCTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGCTCCTGGCCAGGGGCTGCCTCCAGTGGC
CCCCATCCTCAATGCCCACCAGCCCAGCACCCAGCGGGCAGAGATCCACTTCCTGAAGCAAGGCTGTGGTGAAGA
CAAGATCTGCCAGAGCAATCTGCAGCTGGTCCACGCCCGCTTCTGTACCCGGGTCAGCGACACGGAATTCCAACC
TCTGCCCATGGATGTGGATGGAACAACAGCCCTGTTTGCACTGAGTGGGCAGCCAGTCATTGGCCTGGAGCTGAT
GGTCACCAACCTGCCATCGGACCCAGCCCAGCCCCAGGCTGATGGGGATGATGCCCATGAAGCCCAGCTCCTGGT
CATGCTTCCTGACTCACTGCACTACTCAGGGGTCCGGGCCCTGGACCCTGCGGAGAAGCCACTCTGCCTGTCCAA
TGAGAATGCCTCCCATGTTGAGTGTGAGCTGGGGAACCCCATGAAGAGAGGTGCCCAGGTCACCTTCTACCTCAT
CCTTAGCACCTCCGGGATCAGCATTGAGACCACGGAACTGGAGGTAGAGCTGCTGTTGGCCACGATCAGTGAGCA
GGAGCTGCATCCAGTCTCTGCACGAGCCCGTGTCTTCATTGAGCTGCCACTGTCCATTGCAGGAATGGCCATTCC
CCAGCAACTCTTCTTCTCTGGTGTGGTGAGGGGCGAGAGAGCCATGCAGTCTGAGCGGGATGTGGGCAGCAAGGT
CAAGTATGAGGTCACGGTTTCCAACCAAGGCCAGTCGCTCAGAACCCTGGGCTCTGCCTTCCTCAACATCATGTG
GCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACCCAATGCAGGTTGAGCTGGAGGGCGGGCAGGGGCCTGG
GCAGAAAGGGCTTTGCTCTCCCAGGCCCAACATCCTCCACCTGGATGTGGACAGTAGGGATAGGAGGCGGCGGGA
GCTGGAGCCACCTGAGCAGCAGGAGCCTGGTGAGCGGCAGGAGCCCAGCATGTCCTGGTGGCCAGTGTCCTCTGC
TGAGAAGAAGAAAAACATCACCCTGGACTGCGCCCGGGGCACGGCCAACTGTGTGGTGTTCAGCTGCCCACTCTA
CAGCTTTGACCGCGCGGCTGTGCTGCATGTCTGGGCCGTCTCTGGAACAGCACCTTTCTGGAGGAGTACTCAGC
TGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACATCACAGTGAAGTCCTCCATAAAGAACTTGATGCTCCGAGA
TGCCTCCACAGTGATCCCAGTGATGGTATACTTGGACCCCATGGCTGTGGTGGCAGAAGGAGTGCCCTGGTGGGT
CATCCTCCTGGCTGTACTGGCTGGGCTGCTGGTGCTGGCTCTCCTGGTTCTGGTGGTGTCTCCTGGTGGGAGATGGGATTCTT
CAAACGGGCGAAGCACCCCGAGGCCACCGTGCCCCAGTACCATGCGGTGAAGATTCCTCGGGAAGACCGACAGCA
GTTCAAGGAGGAGAAGACGGGCACCATCCTGAGGAACAACTGGGGCAGCCCCCGGGGGAGGGCCCGGATGCACA
CCCCATCCTGGCTGCTGACGGGCATCCCGAGCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCC**TAG**GTTCC
CATGTCCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGT
GGGCTGCTGGTGTCGCATCAAGATTTGGCAGGATCGGCTTCCTCAGGGGCACAGACCTCTCCCACCCACAAGAAC
TCCTCCCACCCAACTTCCCCTTAGAGTGCTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCATGGGGTAGGG
TGAGAAGGGCAGGGGTGTCCTGATGCAAAGGTGGGGAGAAGGGATCCTAATCCCTTCCTCTCCCATTCACCCTGT
GTAACAGGACCCCAAGGACCTGCCTCCCCGGAAGTGCCTTAACCTAGAGGGTCGGGGAGGAGGTTGTGTCACTGA
CTCAGGCTGCTCCTTCTCTAGTTTCCCCTCTCATCTGACCTTAGTTTGCTGCCATCAGTCTAGTGGTTTCGTGGT
TTCGTCTATTTATTAAAAAATATTTGAGAACAAAAAAAAAAAAAAAAAAAA

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

```
MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSWLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMQKESKENQWL
GVSVRSQGPGGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDGGEWKFCEG
RPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGTARVELCAQGSADLAHLDDGPYEA
GGEKEQDPRLIPVPANSYFGFSIDSGKGLVRAEELSFVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGWPDLIVGAPYFFERQEELGGAVYVYLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDLNQDGFPDIAVGAPFDGDGKVFIYHGSSLGVVAKPSQVLE
GEAVGIKSFGYSLSGSLDMDGNQYPDLLVGSLADTAVLFRARPILHVSHEVSIAPRSIDLEQ
PNCAGGHSVCVDLRVCFSYIAVPSSYSPTVALDYVLDADTDRRLRGQVPRVTFLSRNLEEPK
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSYSLQTPRLRRQAPGQGLPPVAP
ILNAHQPSTQRAEIHFLKQGCGEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVDGTTALFA
LSGQPVIGLELMVTNLPSDPAQPQADGDDAHEAQLLVMLPDSLHYSGVRALDPAEKPLCLSN
ENASHVECELGNPMKRGAQVTFYLILSTSGISIETTELEVELLLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKYEVTVSNQGQSLRTLGSAFLNIM
WPHEIANGKWLLYPMQVELEGGQGPGQKGLCSPRPNILHLDVDSRDRRRRELEPPEQQEPGE
RQEPSMSWWPVSSAEKKKNITLDCARGTANCVVFSCPLYSFDRAAVLHVWGRLWNSTFLEEY
SAVKSLEVIVRANITVKSSIKNLMLRDASTVIPVMVYLDPMAVVAEGVPWWVILLAVLAGLL
VLALLVLLLWKMGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILRNNWGSPRREGP
DAHPILAADGHPELGPDGHPGPGTA
```

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCAGA

AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCTGGCT

CACAACAAG**ATG**CTCAAGGTGTCAGCCGTACTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCA

GTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGGGCGGTCGGACGGCGGTAATTTTC

TGGATGATAAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC

AAATTCCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGGAAAACCCTTCGA

TCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTAT

GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG

AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATATTATCCACCTGCAAGCAGTG

CCCAGTGGTCTATCCCAGCCCTGTTTGTGGTTCAGATGGTCATACCTACTCTTTTCAGTGCA

AACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC

CCATGTCCTTCAGATAAGCCCACCAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT

GGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGAA

GTCAAAACAAGAAGACAAAAACATTGCTGAGGCCTGAGAGAAGCAGATTCGATACCAGCATC

TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAACAGACTTGATACAAACTATGACCT

GCTATTGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAATGAACAGTGTACCAAGG

CATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAATATCTAATAATGAGTGGTGCTAC

TGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCA

AGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC

CAACACAATGTCATGGCAGTGTTGGACAGTGCTGGTGTGTTGACAGATATGGAAATGAAGTC

ATGGGATCCAGAATAAATGGTGTTGCAGATTGTGCTATAGATTTTGAGATCTCCGGAGATTT

TGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG

ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGATGAC

CATGATGTATACATT**TGA**TTGATGACAGTTGAAATCAATAAATTCTACATTTCTAATATTTA

CAAAAATGATAGCCTATTTAAAATTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCA

CATATATTTTGTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAACTTTATGTTTAAAT

AAGAATCATTTGCTTTGAGTTTTTATATTCCTTACACAAAAAGAAAATACATATGCAGTCTA

GTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTCACGAGAACAAACTTTGT

AAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAAG

ATAATTCTAAGTGAAATTTAAAATAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGG

AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG

GATAACAGAGAGATACCACATGACTCCAAAAAAAAAAAAAAA

FIGURE 179

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829

><subunit 1 of 1, 436 aa, 1 stop

><MW: 49429, pI: 4.80, NX(S/T): 0

MLKVSAVLCVCAAAWCSQSLAAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR

DEVEDDYFRTWSPGKPFDQALDPAKDPCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA

GVDHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCPCP

SDKPTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTKTLLRPERSRFDTSILPI

CKDSLGWMFNRLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ

RQQDPPCQTELSNIQKRQGVKKLLGQYIPLCDEDGYYKPTQCHGSVGQCWCVDRYGNEVMGS

RINGVADCAIDFEISGDFASGDFHEWTDDEDDEDDIMNDEDEIEDDDEDEGDDDDGGDDHDVYI

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 180

CAGACTCCAGATTTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGACGC
CTCTTTCAGCCCGGGATCGCCCCAGCAGGG**ATG**GGCGACAAGATCTGGCTGCCCTTCCCCGTGCTCCTTCTGGCC
GCTCTGCCTCCGGTGCTGCTGCCTGGGGCGGCCGGCTTCACACCTTCCCTCGATAGCGACTTCACCTTTACCCTT
CCCGCCGGCCAGAAGGAGTGCTTCTACCAGCCCATGCCCCTGAAGGCCTCGCTGGAGATCGAGTACCAAGTTTTA
GATGGAGCAGGATTAGATATTGATTTCCATCTTGCCTCTCCAGAAGGCAAAACCTTAGTTTTTGAACAAAGAAAA
TCAGATGGAGTTCACACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTCAGCACCATT
TCTGAGAAGGTGATTTTCTTTGAATTAATCCTGGATAATATGGGAGAACAGGCACAAGAACAAGAAGATTGGAAG
AAATATATTACTGGCACAGATATATTGGATATGAAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCC
AGACTAAGCAAAAGTGGGCACATACAAATTCTGCTTAGAGCATTTGAAGCTCGTGATCGAAACATACAAGAAAGC
AACTTTGATAGAGTCAATTTCTGGTCTATGGTTAATTTAGTGGTCATGGTGGTGGTGTCAGCCATTCAAGTTTAT
ATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAAGTAGAACT**TAA**AACTCCAAACTAGAGTACGTAACATTGAAA
AATGAGGCATAAAAATGCAATAAACTGTTACAGTCAAGACCATTAATGGTCTTCTCCAAAATATTTTGAGATATA
AAAGTAGGAAACAGGTATAATTTTAATGTGAAAATTAAGTCTTCACTTTCTGTGCAAGTAATCCTGCTGATCCAG
TTGTACTTAAGTGTGTAACAGGAATATTTTGCAGAATATAGGTTTAACTGAATGAAGCCATATTAATAACTGCAT
TTTCCTAACTTTGAAAAATTTTGCAAATGTCTTAGGTGATTTAAATAAATGAGTATTGGGCCTAATTGCAACACC
AGTCTGTTTTTAACAGGTTCTATTACCCAGAACTTTTTTGTAAATGCGGCAGTTACAAATTAACTGTGGAAGTTT
TCAGTTTTAAGTTATAAATCACCTGAGAATTACCTAATGATGGATTGAATAAATCTTTAGACTACAAAAGCCCAA
CTTTTCTCTATTTACATATGCATCTCTCCTATAATGTAAATAGAATAATAGCTTTGAAATACAATTAGGTTTTTG
AGATTTTTATAACCAAATACATTTCAGTGTAACATATTAGCAGAAAGCATTAGTCTTTGTACTTTGCTTACATTC
CCAAAAGCTGACATTTTCACGATTCTTAAAAACACAAAGTTACACTTACTAAAATTAGGACATGTTTTCTCTTTG
AAATGAAGAATATAGTTTAAAAGCTTCCTCCTCCATAGGGACACATTTTCTCTAACCCTTAACTAAAGTGTAGGA
TTTTAAAATTAAATGTGAGGTAAAATAAGTTTATTTTTAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAA
TAATCATGTTATGTTAATTTTAACATGATTGCTGACTTGGATAATTCATTATTACCAGCAGTTATGAAGGAAATA
TTGCTAAAATGATCTGGGCCTACCATAAATAAATATCTCCTTTTCTGAGCTCTAAGAATTATCAGAAAACAGGAA
AGAATTTAGAAAAACTTGAGAAAACCTAATCCAAAATAAAATTCACTTAAGTAGAACTATAAATAAATATCTAGA
ATCTGACTGGCTCATCATGACATCCTACTCATAACATAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG
AAACTTTGGCTGTAGGTTTTTATTTTCTACAAGAATTCTGGTTTGAATTATTTTTGTAAGCAGGTACATTTTATA
AAATGTAAGCCCTACTGTAAGGTTTAGCACTGGGTGTACATATTTATTAAAAATTTTTATTATAACAACTTTTAT
TAAAATGGCCTTTCTGAACACTTTATTTATTGATGTTGAAGTAAGGATTAGAAACATAGACTCCCAAGTTTTAAA
CACCTAAATGTGAATAACCCATATATACAACAAAGTTTCTGCCATCTAGCTTTTTGAAGTCTATGGGGGTCTTAC
TCAAGTACTAGTAATTTAACTTCATCATGAATGAACTATAATTTTTAAGTTATGCCCATTTATAACGTTGTTTAT
GACTACATTGTGAGTTAGAAACAAACTTAAAATTTGGGGTATAGAACCCCTCAACAGGTTAGTAATGCTGGAATT
CTTGATGAGCAATAATGATAACCAGAGAGTGATTTCATTTACACTCATAGTAGTATAAAAAGAGATACATTTCCC
TCTTAGGCCCCTGGGAGAAGAGCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAAAATGAGGTAAATGCCGTAT
ATGATCAATTACCTTAATTGGCCAAGAAAATGCTTCAGGTGTCTAGGGGTATCCTCTGCAACACTTGCAGAACAA
AGGTCAATAAGATCCTTGCCTATGAATACCCCTCCCTTTTGCGCTGTTAAATTTGCAATGAGAAGCAAATTTACA
GTACCATAACTAATAAAGCAGGGTACAGATATAAACTACTGCATCTTTTCTATAAAACTGTGATTAAGAATTCTA
CCTCTCCTGTATGGCTGTTACTGTACTGTACTCTCTGACTCCTTACCTAACAATGAATTTGTTACATAATCTTCT
ACATGTATGATTTGTGCCACTGATCTTAAACCTATGATTCAGTAACTTCTTACCATATAAAAACGATAATTGCTT
TATTTGGAAAAGAATTTAGGAATACTAAGGACAATTATTTTTATAGACAAAGTAAAAAGACAGATATTTAAGAGG
CATAACCAAAAAAGCAAAACTTGTAAACAGAGTAAAAATCTTTAATATTTCTAAAGACATACTGTTTATCTGCTT
CATATGCTTTTTTTAATTTCACTATTCCATTTCTAAATTAAAGTTATGCTAAATTGAGTAAGCTGTTTATCACTT
AACAGCTCATTTTGTCTTTTTCAATATACAAATTTTAAAAATACTACAATATTTAACTAAGGCCCAACCGATTTC
CATAATGTAGCAGTTACCGTGTTCACCTCACACTAAGGCCTAGAGTTTGCTCTGATATGCATTTGGATGATTAAT
GTTATGCTGTTCTTTCATGTGAATGTCAAGACATGGAGGGTGTTTGTAATTTTATGGTAAAATTAATCCTTCTTA
CACATAATGGTGTCTTAAAATTGACAAAAAATGAGCACTTACAATTGTATGTCTCCTCAAATGAAGATTCTTTAT
GTGAAATTTTAAAAGACATTGATTCCGCATGTAAGGATTTTTCATCTGAAGTACAATAATGCACAATCAGTGTTG
CTCAAACTGCTTTATACTTATAAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATATGTATATAATAA
AAATTATCAAAGGAAAA

FIGURE 181

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196

><subunit 1 of 1, 229 aa, 1 stop

><MW: 26017, pI: 4.73, NX(S/T): 0

MGDKIWLPFPVLLLAALPPVLLPGAAGFTPSLDSDFTFTLPAGQKECFYQPMPLKASLEIEY

QVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCFDNTFSTISEKVIFFEL

ILDNMGEQAQEQEDWKKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN

IQESNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

FIGURE 182

CCATCCCTGAGATCTTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCAGAT
CTCACCAGAGAGTCGCAGACACT**ATG**CTGCCTCCCATGGCCCTGCCCAGTGTGTCCTGGATG
CTGCTTTCCTGCCTCATTCTCCTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGAACTGCC
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT
TTTTGTCACCAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCCTCCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACATCTGGATTGGGCTCCATGACCCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCC
TCCACCATCTTAAACCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCTATGTCTGCAAGTTCAAGGAC**TAG**GGCAGGT
GGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCCTGGAAGAGAATATTCTCCCCAAACTGCCCTACCTGACTACCTTGTCATGATCCTCC
TTCTTTTTCCTTTTTCTTCACCTTCATTTCAGGCTTTTCTCTGTCTTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

FIGURE 183

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965

<subunit 1 of 1, 175 aa, 1 stop

<MW: 19330, pI: 7.25, NX(S/T): 1

MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM

DADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGWEWSS

TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAKLPYVCKFKD

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

FIGURE 184

CCAGTCTGTCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACC**ATG**CGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGCTGGC
TGCCTGCGGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCCACAGGAGTGT
CGGACTGTGTCACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCGGGAGATAGTGTACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCTGCCCGTGTCCTGCTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCCGCTCTGAACAGCCTCCACTGCGGGGCCCTCACGCTC
CTCCCACTCTTGAGCCTCCGACTG**TAG**AGTCCCCGCCCACCCCCATGGCCCTATGCGGCCCA
GCCCCGAATGCCTTGAAGAAGTGCCCCCTGCACCAGGAAAAAAAAAAAAAAAA

FIGURE 185

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405

<subunit 1 of 1, 125 aa, 1 stop

<MW: 13115, pI: 5.90, NX(S/T): 1

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP

FQGDSTVTKSCASKCKPSDVDGIGQTLPVSCCNTELCNVDGAPALNSLHCGALTLLPLLSLRL

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC

ACGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTTCTCCCTCTT

GAGTCCTTCTGAG**ATG**ATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTCGCGATGG

TAGCGGCGGCTCTCGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAACTCGGTTCTC

AATTCCAACGCTATCAAGAACCTGCCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTC

TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA

ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT

CCCACCCGCGGAGGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACG

CTGCATGCGTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTT

CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT

CATAGCACCTTGGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAA

AGGACAAGAAGGTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA

GACACTTCTGGTCCAAGATCTGTAAACCTGTCCTGAAAGAAGGTCAAGTGTGTACCAAGCAT

AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTC

TTGCCGGATACAGAAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTCAGA

GACAC**TAA**ACCAGCTATCCAAATGCAGTGAACTCCTTTTATATAATAGATGCTATGAAAACC

TTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCAT

TCCAATAAC₄CCTTCCAAAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACTCCCCTG

TGATTGCAGTAAATTACTGTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAATGCAATGA

AACTTTTAATTATTTTTCTAAAGGTGCTGCACTGCCTATTTTTCCTCTTGTTATGTAAATTT

TTGTACACATTGATTGTTATCTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATT

TCAGCTTATAGTTCTTAAAAGCATAACCCTTTACCCCATTTAATTCTAGAGTCTAGAACGCA

AGGATCTCTTGGAATGACAAATGATAGGTACCTAAAATGTAACATGAAAATACTAGCTTATT

TTCTGAAATGTACTATCTTAATGCTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTGA

AATAAAATTTAACATTTAAAAAAAAAAAAA

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530

<subunit 1 of 1, 266 aa, 1 stop

<MW: 28672, pI: 8.85, NX(S/T): 1

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA

APGILYPGGNKYQTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRKRCMRH

AMCCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG

SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQ

KDHHQASNSSRLHTCQRH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

FIGURE 188

TGTGTTTCCCTGCAGTCAGAATTTGGGACNGCAGGGGTTCCCGGACCTGATTTTGCAGCGGA

ACGGGAAGGTTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTTTTCTTTCTCCTTCNG

GAGTCCTTNTGAGANGATGGTTTTGGGCGCAGCGGGAGCTAACCCGGTTTTTTGTNGCGATG

GTAGCGGCGGTTTTCGGCGGCCACCTTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTTC

AATTCCAACGNTATCAAGAACCTGCCCCCACCGNTGGGCGGCGCTGCGGGGCACCCAGGNTT

TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA

ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT

CCCACCCGCGGAGGGGANGCGGGCGTGCAAATNTGTNTNGCCTGCAGGAAGCGCCGAAAACG

CTGCATGCGTCANGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTNTT

CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT

CATAGCACCTTGGATGGG

FIGURE 189

GAGGAACCTACCGGTACCGGCCGCGCGCTGGTAGTCGCCGGTGTGGCTGCACCTCACCAATCCCGTGCGCCGCGG
CTGGGCCGTCGGAGAGTGCGTGTGCTTCTCTCCTGCACGCGGTGCTTGGGCTCGGCCAGGCGGGGTCCGCCGCCA
GGGTTTGAGGATGGGGGAGTAGCTACAGGAAGCGACCCCGCGATGGCAAGGTATATTTTTGTGGAATGAAAAGGA
AGTATTAGAAATGAGCTGAAGACCATTCACAGATTAATATTTTTGGGGACAGATTTGTGATGCTTGATTCACCCT
TGAAGTAATGTAGACAGAAGTTCTCAAATTTGCATATTACATCAACTGGAACCAGCAGTGAATCTTAATGTTCAC
TTAAATCAGAACTTGCATAAGAAAGAGA**ATG**GGAGTCTGGTTAAATAAAGATGACTATATCAGAGACTTGAAAAG
GATCATTCTCTGTTTTCTGATAGTGTATATGGCCATTTTAGTGGGCACAGATCAGGATTTTTACAGTTTACTTGG
AGTGTCCAAAACTGCAAGCAGTAGAGAAATAAGACAAGCTTTCAAGAAATTGGCATTGAAGTTACATCCTGATAA
AAACCCGAATAACCCAAATGCACATGGCGATTTTTTAAAAATAAATAGAGCATATGAAGTACTCAAAGATGAAGA
TCTACGGAAAAAGTATGACAAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGGCCAGTATGAAAGCTGGAA
CTATTATCGTTATGATTTTGGTATTTATGATGATGATCCTGAAATCATAACATTGGAAAGAAGAGAATTTGATGC
TGCTGTTAATTCTGGAGAACTGTGGTTTGTAAATTTTTACTCCCCAGGCTGTTCACACTGCCATGATTTAGCTCC
CACATGGAGAGACTTTGCTAAAGAAGTGGATGGGTTACTTCGAATTGGAGCTGTTAACTGTGGTGATGATAGAAT
GCTTTGCCGAATGAAAGGAGTCAACAGCTATCCCAGTCTCTTCATTTTTCGGTCTGGAATGGCCCCAGTGAAATA
TCATGGAGACAGATCAAAGGAGAGTTTAGTGAGTTTTGCAATGCAGCATGTTAGAAGTACAGTGACAGAACTTTG
GACAGGAAATTTTGTCAACTCCATACAAACTGCTTTTGCTGCTGGTATTGGCTGGCTGATCACTTTTTGTTCAAA
AGGAGGAGATTGTTTGACTTCACAGACACGACTCAGGCTTAGTGGCATGTTGTTTCTCAACTCATTGGATGCTAA
AGAAATATATTTGGAAGTAATACATAATCTTCCAGATTTTGAACTACTTTCGGCAAACACACTAGAGGATCGTTT
GGCTCATCATCGGTGGCTGTTATTTTTTCATTTTGGAAAAAATGAAAATTCAAATGATCCTGAGCTGAAAAAACT
AAAAACTCTACTTAAAAATGATCATATTCAAGTTGGCAGGTTTGACTGTTCCTCTGCACCAGACATCTGTAGTAA
TCTGTATGTTTTTCAGCCGTCTCTAGCAGTATTTAAAGGACAAGGAACCAAAGAATATGAAATTCATCATGGAAA
GAAGATTCTATATGATATACTTGCCTTTGCCAAAGAAAGTGTGAATTCTCATGTTACCACGCTTGGACCTCAAAA
TTTTCCTGCCAATGACAAAGAACCATGGCTTGTTGATTTCTTTGCCCCCTGGTGTCCACCATGTCGAGCTTTACT
ACCAGAGTTACGAAGAGCATCAAATCTTCTTTATGGTCAGCTTAAGTTTGGTACACTAGATTGTACAGTTCATGA
GGGACTCTGTAACATGTATAACATTCAGGCTTATCCAACAACAGTGGTATTCAACCAGTCCAACATTCATGAGTA
TGAAGGACATCACTCTGCTGAACAAATCTTGGAGTTCATAGAGGATCTTATGAATCCTTCAGTGGTCTCCCTTAC
ACCCACCACCTTCAACGAACTAGTTACACAAAGAAAACACAACGAAGTCTGGATGGTTGATTTCTATTCTCCGTG
GTGTCATCCTTGCCAAGTCTTAATGCCAGAATGGAAAAGAAATGGCCCCGGACATTAACTGGACTGATCAACGTGGG
CAGTATAGATTGCCAACAGTATCATTCTTTTTGTGCCCAGGAAACCGTTCAAAGATACCCTGAGATAAGATTTTT
TCCCCCAAAATCAAATAAAGCTTATCAGTATCACAGTTACAATGGTTGGAATAGGGATGCTTATTCCCTGAGAAT
CTGGGGTCTAGGATTTTTACCTCAAGTATCCACAGATCTAACACCTCAGACTTTCAGTGAAAAAGTTCTACAAGG
GAAAAATCATTGGGTGATTGATTTCTATGCTCCTTGGTGTGGACCTTGCCAGAATTTTGCTCCAGAATTTGAGCT
CTTGGCTAGGATGATTAAAGGAAAAGTGAAAGCTGGAAAAGTAGACTGTCAGGCTTATGCTCAGACATGCCAGAA
AGCTGGGATCAGGGCCTATCCAACTGTTAAGTTTTATTTCTACGAAAGAGCAAAGAGAAATTTTCAAGAAGAGCA
GATAAATACCAGAGATGCAAAAGCAATCGCTGCCTTAATAAGTGAAAAATTGGAAACTCTCCGAAATCAAGGCAA
GAGGAATAAGGATGAACTT**TGA**TAATGTTGAAGATGAAGAAAAAGTTTAAAAGAAATTCTGACAGATGACATCAG
AAGACACCTATTTAGAATGTTACATTTATGATGGGAATGAATGAACATTATCTTAGACTTGCAGTTGTACTGCCA
GAATTATCTACAGCACTGGTGTAAAAGAAGGGTCTGCAAACTTTTTCTGTAAAGGGCCGGTTTATAAATATTTTA
GACTTTGCAGGCTATAATATATGGTTCACACATGAGAACAAGAATAGAGTCATCATGTATTCTTTGTTATTTGCT
TTTAACAACCTTTAAAAAATATTAAAACGATTCTTAGCTCAGAGCCATACAAAAGTAGGCTGGATTCAGTCCATG
GACCATAGATTGCTGTCCCCCTCGACGGACTTATAATGTTTCAGGTGGCTGGCTTGAACATGAGTCTGCTGTGCT
ATCTACATAAATGTCTAAGTTGTATAAAGTCCACTTTCCCTTCACGTTTTTTGGCTGACCTGAAAAGAGGTAACT
TAGTTTTTGGTCACTTGTTCTCCTAAAAATGCTATCCCTAACCATATATTTATATTTCGTTTTAAAAACACCCAT
GATGTGGCACAGTAAACAAACCCTGTTATGCTGTATTATTATGAGGAGATTCTTCATTGTTTTCTTTCCTTCTCA
AAGGTTGAAAAAATGCTTTTAATTTTTCACAGCCGAGAAACAGTGCAGCAGTATATGTGCACACAGTAAGTACAC
AAATTTGAGCAACAGTAAGTGCACAAATTCTGTAGTTTGCTGTATCATCCAGGAAAACCTGAGGGAAAAAAATTA
TAGCAATTAACTGGGCATTGTAGAGTATCCTAAATATGTTATCAAGTATTTAGAGTTCTATATTTTAAAGATATA
TGTGTTCATGTATTTTCTGAAATTGCTTTCATAGAAATTTTCCCACTGATAGTTGATTTTTGAGGCATCTAATAT
TTACATATTTGCCTTCTGAACTTTGTTTTGACCTGTATCCTTTATTTACATTGGGTTTTTCTTTCATAGTTTTTGG
TTTTTCACTCCTGTCCAGTCTATTTATTATTCAAATAGGAAAAATTACTTTACAGGTTGTTTTACTGTAGCTTAT
AATGATACTGTAGTTATTCCAGTTACTAGTTTACTGTCAGAGGGCTGCCTTTTTCAGATAAATATTGACATAATA
ACTGAAGTTATTTTTATAAGAAAATCAAGTATATAAATCTAGGAAAGGGATCTTCTAGTTTCTGTGTTGTTTAGA
CTCAAAGAATCACAAATTTGTCAGTAACATGTAGTTGTTTAGTTATAATTCAGAGTGTACAGAATGGTAAAAATT
CCAATCAGTCAAAAGAGGTCAATGAATTAAAAGGCTTGCAACTTTTTCAAAAAAAAAAAAAAAAAA

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439

<subunit 1 of 1, 747 aa, 1 stop

<MW: 86127, pI: 7.46, NX(S/T): 2

MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTDQDFYSLLGVSKTASSREIRQAFKKLALKL

HPDKNPNNPNAHGDFLKINRAYEVLKDEDLRKKYDKYGEKGLEDNQGGQYESWNYYRYDFGI

YDDDPEIITLERREFDAAVNSGELWFVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC

GDDRMLCRMKGVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNS

IQTAFAAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN

TLEDRLAHHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICSNLYVFQP

SLAVFKGQGTKEYEIHHGKKILYDILAFAKESVNSHVTTLGPQNFPANDKEPWLVDFFAPWC

PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIHEYEGHHS

AEQILEFIEDLMNPSVVSLTPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART

LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLG

FLPQVSTDLTPQTFSEKVLQGKNHWVIDFYAPWCGPCQNFAPEFELLARMIKGKVKAGKVDC

QAYAQTCQKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAIAALISEKLETLRNQGKRNKDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAA

GCC**ATG**AACATCATCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGA

GTCGTTGGTGAAGTTTTTCATTCCTCAGAGGAGAAAATCTGTGGCTGGGGAGATTGTTCTCA

TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTTGCAAAACGACAGAGC

ATATTGGTTCTGTGGGATATTAATAAGCGCGGTGTGGAGGAAACTGCAGCTGAGTGCCGAAA

ACTAGGCGTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT

CTCTAAATCAGGTGAAGAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA

GTATATCCAGCCGATCTTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAA

CATCCTAGGACATTTTTGGATCACAAAAGCACTTCTTCCATCGATGATGGAGAGAAATCATG

GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCCTTACCTCATCCCATAT

TGTTCCAGCAAATTTGCCGCTGTTGGCTTTCACAGAGGTCTGACATCAGAACTTCAGGCCTT

GGGAAAAACTGGTATCAAAACCTCATGTCTCTGCCCAGTTTTTGTGAATACTGGGTTCACCA

AAAATCCAAGCACAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA

GATGGAATACTTACCAATAAGAAAATGATTTTTGTTCCATCGTATATCAATATCTTTCTGAG

ACTACAGAAGTTTCTTCCTGAACGCGCCTCAGCGATTTTAAATCGTATGCAGAATATTCAAT

TTGAAGCAGTGGTTGGCCACAAAATCAAAATGAAA**TGA**ATAAATAAGCTCCAGCCAGAGATG

TATGCATGATAATGATATGAATAGTTTCGAATCAATGCTGCAAAGCTTTATTTCACATTTTT

TCAGTCCTGATAATATTAAAAACATTGGTTTGGCACTAGCAGCAGTCAAACGAACAAGATTA

ATTACCTGTCTTCCTGTTTCTCAAGAATATTTACGTAGTTTTTCATAGGTCTGTTTTTCCTT

TCATGCCTCTTAAAAACTTCTGTGCTTACATAAACATACTTAAAAGGTTTTCTTTAAGATAT

TTTATTTTTCCATTTAAAGGTGGACAAAGCTACCTCCCTAAAAGTAAATACAAAGAGAACT

TATTTACACAGGGAAGGTTTAAGACTGTTCAAGTAGCATTCCAATCTGTAGCCATGCCACAG

AATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTTT

ATCTCAACCTGGACATATTTTAAGATTCAGCATTTGAAAGATTTCCCTAGCCTCTTCCTTTT

TCATTAGCCCAAAACGGTGCAACTCTATTCTGGACTTTATTACTTGATTCTGTCTTCTGTAT

AACTCTGAAGTCCACCAAAAGTGGACCCTCTATATTTCCTCCCTTTTTATAGTCTTATAAGA

TACATTATGAAAGGTGACCGACTCTATTTTAAATCTCAGAATTTTAAGTTCTAGCCCCATGA

TAACCTTTTTCTTTGTAATTTATGCTTTCATATATCCTTGGTCCCAGAGATGTTTAGACAAT

TTTAGGCTCAAAAATTAAAGCTAACACAGGAAAAGGAACTGTACTGGCTATTACATAAGAAA

CAATGGACCCAAGAGAAGAA

FIGURE 192

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409

<subunit 1 of 1, 300 aa, 1 stop

<MW: 33655, pI: 9.31, NX(S/T): 1

MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSI
LVLWDINKRGVEETAAECRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDVTIVVNNAGTV
YPADLLSTKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYC
SSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFVNTGFTKNPSTRLWPVLETDEVVRSLID
GILTNKKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 193

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCCAGCAGG

ATGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGGAGAGGGCCCAGCCCGCCCGGGGC

AGG**ATG**ACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTCATGATCCT

GCTGATCATCGTGTACTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGCACACGTCCTTCT

CTAGGCCGCACACGGGGCCGCCGCTGCCCACGCCCGGGCCGGACAGGGACAGGGAGCTCACG

GCCGACTCCGATGTCGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGA

CCTTCCCAGAAAGGAGACGGAGCAGCCGCCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAG

GCTACGACTGGTCCCCGCGCGACGCCCGGCGCAGCCCAGACCAGGGCCGGCAGCAGGCGGAG

CGGAGGAGCGTGCTGCGGGGCTTCTGCGCCAACTCCAGCCTGGCCTTCCCCACCAAGGAGCG

CGCATTCGACGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGGACGACCGGCACGGGG

CCATCTACTGCTACGTGCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTG

AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCA

CGTGCACAACGCCAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCT

CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCCTCTTCGTGCGCGACCCC

TTCGTGCGCCTGATCTCCGCCTTCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCG

CAAGTTCGCCGTGCCCATGCTGCGGCTGTACGCCAACCACACCAGCCTGCCCGCCTCGGCGC

GCGAGGCCTTCCGCGCTGGCCTCAAGGTGTCCTTCGCCAACTTCATCCAGTACCTGCTGGAC

CCGCACACGGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCA

CCCGTGCCAGATCGACTACGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGC

AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTTCCCCCCGAGCTACCGGAACAGG

ACCGCCAGCAGCTGGGAGGAGGACTGGTTCGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCT

GTATAAACTCTACGAGGCCGACTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCC

GAGAC**TGA**AAGCTTTCGCGTTGCTTTTTCTCGCGTGCCTGGAACCTGACGCACGCGCACTCC

AGTTTTTTTATGACCTACGATTTTGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCC

ATTGAGTACTGTATCGATATTGTTTTTTAAGATTAATATATTTCAGGTATTTAATACGA

FIGURE 194

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112

<subunit 1 of 1, 414 aa, 1 stop

<MW: 48414, pI: 9.54, NX(S/T): 4

MTKARLFRLWLVLGSVFMILLIIVYWDSAGAAHFYLHTSFSRPHTGPPLPTPGPDRDRELTA

DSDVDEFLDKFLSAGVKQSDLPRKETEQPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAER

RSVLRGFCANSSLAFPTKERAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRVMIVLS

GSLLHRGAPYRDPLRIPREHVHNASAHLTFNKFWRRYGKLSRHLMKVKLKKYTKFLFVRDPF

VRLISAFRSKFELENEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDP

HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLETLDEDAAQLLQLLQVDRQLRFPPSYRNRT

ASSWEEDWFAKIPLAWRQQLYKLYEADFVLFGYPKPENLLRD

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

FIGURE 195

TCGGGCCAGAATTCGGCACGAGGCGGCACGAGGGCGACGGCCTCACGGGGCTTTGGAGGTGA

AAGAGGCCCAGAGTAGAGAGAGAGAGAGACCGACGTACACGGG**ATG**GCTACGGGAACGCGCT

ATGCCGGGAAGGTGGTGGTCGTGACCGGGGGCGGGCGCGGCATCGGAGCTGGGATCGTGCGC

GCCTTCGTGAACAGCGGGGCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGCCGGGC

CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTTATCCTCTGTGATGTGACTCAGGAAGATGATG

TGAAGACCCTGGTTTCTGAGACCATCCGCCGATTTGGCCGCCTGGATTGTGTTGTCAACAAC

GCTGGCCACCACCCACCCCCACAGAGGCCTGAGGAGACCTCTGCCCAGGGATTCCGCCAGCT

GCTGGAGCTGAACCTACTGGGGACGTACACCTTGACCAAGCTCGCCCTCCCCTACCTGCGGA

AGAGTCAAGGGAATGTCATCAACATCTCCAGCCTGGTGGGGGCAATCGGCCAGGCCCAGGCA

GTTCCCTATGTGGCCACCAAGGGGGCAGTAACAGCCATGACCAAAGCTTTGGCCCTGGATGA

AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCCAGGAAACATCTGGACCCCGCTGTGGG

AGGAGCTGGCAGCCTTAATGCCAGACCCTAGGGCCACAATCCGAGAGGGCATGCTGGCCCAG

CCACTGGGCCGCATGGGCCAGCCCGCTGAGGTCGGGGCTGCGGCAGTGTTCCTGGCCTCCGA

AGCCAACTTCTGCACGGGCATTGAACTGCTCGTGACGGGGGGTGCAGAGCTGGGGTACGGGT

GCAAGGCCAGTCGGAGCACCCCCGTGGACGCCCCCGATATCCCTTCC**TGA**TTTCTCTCATTT

CTACTTGGGGCCCCCTTCCTAGGACTCTCCCACCCCAAACTCCAACCTGTATCAGATGCAGC

CCCCAAGCCCTTAGACTCTAAGCCCAGTTAGCAAGGTGCCGGGTCACCCTGCAGGTTCCCAT

AAAAACGATTTGCAGCC

FIGURE 196

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045

<subunit 1 of 1, 270 aa, 1 stop

<MW: 28317, pI: 6.00, NX(S/T): 1

MATGTRYAGKVVVVTGGGRGIGAGIVRAFVNSGARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRFGRLDCVVNNAGHHPPPQRPEETSAQGFRQLLELNLLGTYTLTKL
ALPYLRKSQGNVINISSLVGAIGQAQAVPYVATKGAVTAMTKALALDESPYGVRVNCISPGN
IWTPLWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAAVFLASEANFCTGIELLVTGG
AELGYGCKASRSTPVDAPDIPS

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

FIGURE 197

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGA**ATG**GACTGGCCTCACAACCTG

CTGTTTCTTCTTACCATTTCCATCTTCCTGGGGCTGGGCCAGCCCAGGAGCCCCAAAAGCAA

GAGGAAGGGGCAAGGGCGGCCTGGGCCCCTGGCCCCTGGCCCTCACCAGGTGCCACTGGACC

TGGTGTCACGGATGAAACCGTATGCCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATG

GTGGCCCAGCTGAGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTTGCAGCT

GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCCAGCC

GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGGCTGTGTGAACCCCTTCACC

ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTCAGCCAGGTTCCTGTGCGCCGCCG

CCTCTGCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG

CTGTGGGCTGCACCTGCATCTTC**TGA**ATCACCTGGCCCAGAAGCCAGGCCAGCAGCCCGAGA

CCATCCTCCTTGCACCTTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAA

GCAAG

FIGURE 198

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294

<subunit 1 of 1, 180 aa, 1 stop

<MW: 20437, pI: 9.58, NX(S/T): 1

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEEY

ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPSRIPVDLPEARCLCL

GCVNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPPRTGPCRQRAVMETIAVGCTCIF

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

FIGURE 199

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCCTTGAAAAACCCGGCGGG
CGAGCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCACACTCCCCGCCGAGAAGCCTCGCTCG
GCGCCCAAC**ATG**GCGGGTGGGCGCTGCGGCCCGCAGCTAACGGCGCTCCTGGCCGCCTGGAT
CGCGGCTGTGGCGGCGACGGCAGGCCCCGAGGAGGCCGCGCTGCCGCCGGAGCAGAGCCGGG
TCCAGCCCATGACCGCCTCCAACTGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTT
TACGCCCCATGGTGTCCATCCTGCCAGCAGACTGATTCAGAATGGGAGGCTTTTGCAAAGAA
TGGTGAAATACTTCAGATCAGTGTGGGGAAGGTAGATGTCATTCAAGAACCAGGTTTGAGTG
GCCGCTTCTTTGTCACCACTCTCCCAGCATTTTTTCATGCAAAGGATGGGATATTCCGCCGT
TATCGTGGCCCAGGAATCTTCGAAGACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATC
AGTCGAGCCTCTGACTGGCTGGAAATCCCCAGCTTCTCTAACGATGTCTGGAATGGCTGGTC
TTTTTAGCATCTCTGGCAAGATATGGCATCTTCACAACTATTTCACAGTGACTCTTGGAATT
CCTGCTTGGTGTTCTTATGTGTTTTTCGTCATAGCCACCTTGGTTTTTGGCCTTTTTATGGG
TCTGGTCTTGGTGGTAATATCAGAATGTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGC
GTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCGGAG
GAGGAAAAAGATGATTCAAATGAAGAAGAAAACAAAGACAGCCTTGTAGATGATGAAGAAGA
GAAAGAAGATCTTGGCGATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGGACAACTTGGCTG
CTGGTGTGGATGAGGAGAGAAGTGAGGCCAATGATCAGGGGCCCCCAGGAGAGGACGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCCTGCCCAGC
TGACACAGAGGTGGTGGAAGACTCCTTGAGGCAGCGTAAAAGTCAGCATGCTGACAAGGGAC
TG**TAG**ATTTAATGATGCGTTTTCAAGAATACACACCAAAACAATATGTCAGCTTCCCTTTGG
CCTGCAGTTTGTACCAAATCCTTAATTTTTCCTGAATGAGCAAGCTTCTCTTAAAAGATGCT
CTCTAGTCATTTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGT
GACAATCAGGATATAGAAAAACAAACGTAGTGTTGGGATCTGTTTGGAGACTGGGATGGGAA
CAAGTTCATTTACTTAGGGGTCAGAGAGTCTCGACCAGAGGAGGCCATTCCCAGTCCTAATC
AGCACCTTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCT
CCTGAGCATCCCCAAAGTGTAACGTAGAAGCCTTGCATCCTTTTCTTGTGTAAAGTATTTAT
TTTTGTCAAATTGCAGGAAACATCAGGCACCACAGTGCATGAAAAATCTTTCACAGCTAGAA
ATTGAAAGGGCCTTGGGTATAGAGAGCAGCTCAGAAGTCATCCCAGCCCTCTGAATCTCCTG
TGCTATGTTTTATTTCTTACCTTTAATTTTTCCAGCATTTCCACCATGGGCATTCAGGCTCT
CCACACTCTTCACTATTATCTCTTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAAC
TGTGTTTGTTCATTCTGACCTAAGGGGTTTAGATAATCAGTAACCATAACCCCTGAAGCTGT
GACTGCCAAACATCTCAAATGAAATGTTGTGGCCATCAGAGACTCAAAAGGAAGTAAGGATT
TTACAAGACAGATTAAAAAAAAATTGTTTTGTCCAAAATATAGTTGTTGTTGATTTTTTTTT
AAGTTTTCTAAGCAATATTTTTCAAGCCAGAAGTCCTCTAAGTCTTGCCAGTACAAGGTAGT
CTTGTGAAGAAAGTTGAATACTGTTTTGTTTTCATCTCAAGGGGTTCCCTGGGTCTTGAAC
TACTTTAATAATAACTAAAAAACCACTTCTGATTTTCCTTCAGTGATGTGCTTTTGGTGAAA
GAATTAATGAACTCCAGTACCTGAAAGTGAAAGATTTGATTTTGTTTCCATCTTCTGTAATC
TTCCAAGAATTATATCTTTGTAAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTTCTTT

FIGURE 200

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433

<subunit 1 of 1, 349 aa, 1 stop

<MW: 38952, pI: 4.34, NX(S/T): 1

MAGGRCGPQLTALLAAWIAAVAATAGPEEAALPPEQSRVQPMTASNWTLVMEGEWMLKFYAP

WCPSCQQTDSEWEAFAKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFRRYRG

PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW

CSYVFFVIATLVFGLFMGLVLVVISECFYVPLPRHLSERSEQNRRSEEAHRAEQLQDAEEEK

DDSNEEENKDSLVDDEEEKEDLGDEDEAEEEEEDNLAAGVDEERSEANDQGPPGEDGVTRE

EVEPEEAEEGISEQPCPADTEVVEDSLRQRKSQHADKGL

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins.   (homologous region to disulfide
isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

FIGURE 201

```
ATCTGGTTGAACTACTTAAGCTTAATTTGTTAAACTCCGGTAAGTACCTAGCCCACATGATT
TGACTCAGAGATTCTCTTTTGTCCACAGACAGTCATCTCAGGGGCAGAAAGAAAAGAGCTCC
CAAATGCTATATCTATTCAGGGGCTCTCAAGAACAATGGAATATCATCCTGATTTAGAAAAT
TTGGATGAAGATGGATATACTCAATTACACTTCGACTCTCAAAGCAATACCAGGATAGCTGT
TGTTTCAGAGAAGGATCGTGTGCTGCATCTCCTCCTTGGCGCCTCATTGCTGTAATTTTGG
GAATCCTATGCTTGGTAATACTGGTGATAGCTGTGGTCCTGGGTACCATGGGGGTTCTTTCC
AGCCCTTGTCCTCCTAATTGGATTATATATGAGAAGAGCTGTTATCTATTCAGCATGTCACT
AAATTCCTGGGATGGAAGTAAAAGACAATGCTGGCAACTGGGCTCTAATCTCCTAAAGATAG
ACAGCTCAAATGAATTGGGATTTATAGTAAAACAAGTGTCTTCCCAACCTGATAATTCATTT
TGGATAGGCCTTTCTCGGCCCCAGACTGAGGTACCATGGCTCTGGGAGGATGGATCAACATT
CTCTTCTAACTTATTTCAGATCAGAACCACAGCTACCCAAGAAAACCCATCTCCAAATTGTG
TATGGATTCACGTGTCAGTCATTTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTTGT
GAGAAGAAGTTTTCAATGTAAGAGGAAGGGTGGAGAAGGAGAGAGAAATATGTGAGGTAGTA
AGGAGGACAGAAAACAGAACAGAAAGAGTAACAGCTGAGGTCAAGATAAATGCAGAAAATG
TTTAGAGAGCTTGGCCAACTGTAATCTTAACCAAGAAATTGAAGGGAGAGGCTGTGATTTCT
GTATTTGTCGACCTACAGGTAGGCTAGTATTATTTTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGGCAGCCAAGCTTGAGTTTTTATTTTTTATTTATTTATTTTTTTGAGATAGGGTCT
CACTTTGTTACCCAGGCTGGAGTGCAGTGGCACAATCTCGACTCACTGCAGCTATCTCTCGC
CTCAGCCCCTCAAGTAGCTGGGACTACAGGTGCATGCCACCATGCCAGGCTAATTTTTGGTG
TTTTTTGTAGAGACTGGGTTTTGCCATGTTGACCAAGCTGGTCTCTAACTCCTGGGCTTAAG
TGATCTGCCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGCCACCACACCTGGC
CCCAAGCTTGAATTTTCATTCTGCCATTGACTTGGCATTTACCTTGGGTAAGCCATAAGCGA
ATCTTAATTTCTGGCTCTATCAGAGTTGTTTCATGCTCAACAATGCCATTGAAGTGCACGGT
GTGTTGCCACGATTTGACCCTCAACTTCTAGCAGTATATCAGTTATGAACTGAGGGTGAAAT
ATATTTCTGAATAGCTAAATGAAGAAATGGGAAAAAATCTTCACCACAGTCAGAGCAATTTT
ATTATTTTCATCAGTATGATCATAATTATGATTATCATCTTAGTAAAAAGCAGGAACTCCTA
CTTTTTCTTTATCAATTAAATAGCTCAGAGAGTACATCTGCCATATCTCTAATAGAATCTTT
TTTTTTTTTTTTTTTTTTTGAGACAGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAACGG
CACGATCTCGGCTCACCGCAACCTCCGCCCCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCT
CCCAAGTAGCTGGGATTACAGTCAGGCACCACCACACCCGGCTAATTTTGTATTTTTTTAGT
AGAGACAGGGTTTCTCCATGTCGGTCAGGGTAGTCCCGAACTCCTGACCTCAAGTGATCTGC
CTGCCTCGGCCTCCCAAGTGCTGGGATTACAGGCGTGAGCCACTGCACCCAGCCTAGAATCT
TGTATAATATGTAATTGTAGGGAAACTGCTCTCATAGGAAAGTTTTCTGCTTTTTAAATACA
AAAATACATAAAAATACATAAAATCTGATGATGAATATAAAAAAGTAACCAACCTCATTGGA
ACAAGTATTAACATTTTGGAATATGTTTTATTAGTTTTGTGATGTACTGTTTTACAATTTTT
ACCATTTTTTTCAGTAATTACTGTAAAATGGTATTATTGGAATGAAACTATATTTCCTCATG
TGCTGATTTGTCTTATTTTTTTCATACTTTCCCACTGGTGCTATTTTTATTTCCAATGGATA
TTTCTGTATTACTAGGGAGGCATTTACAGTCCTCTAATGTTGATTAATATGTGAAAGAAAT
TGTACCAATTTTACTAAATTATGCAGTTTAAAATGGATGATTTTATGTTATGTGGATTTCAT
TTCAATAAAAAAAAACTCTTATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 202

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912
<subunit 1 of 1, 201 aa, 1 stop
<MW: 22563, pI: 4.87, NX(S/T): 1
MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSCAASPPWRLIAVILGILCLVILVIAV
VLGTMGVLSSPCPPNWIIYEKSCYLFSMSLNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM
```

Important features:

Type II transmembrane domain:
amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 197-200

N-myristoylation sites.
amino acids 35-40 and 151-156

Homologous region to LDL receptor
amino acids 34-67 and 70-200.

FIGURE 203

GGAAGGGGAGGAGCAGGCCACACAGGCACAGGCCGGTGAGGGACCTGCCCAGACCTGGAGGGTCTCGCTCTGTCA
CACAGGCTGGAGTGCAGTGGTGTGATCTTGGCTCATCGTAACCTCCACCTCCCGGGTTCAAGTGATTCTCATGCC
TCAGCCTCCCGAGTAGCTGGGATTACAGGTGGTGACTTCCAAGAGTGACTCCGTCGGAGGAAA**ATG**ACTCCCCAG
TCGCTGCTGCAGACGACACTGTTCCTGCTGAGTCTGCTCTTCCTGGTCCAAGGTGCCCACGGCAGGGGCCACAGG
GAAGACTTTCGCTTCTGCAGCCAGCGGAACCAGACACACAGGAGCAGCCTCCACTACAAACCCACACCAGACCTG
CGCATCTCCATCGAGAACTCCGAAGAGGCCCTCACAGTCCATGCCCCTTTCCCTGCAGCCCCACCCTGCTTCCCGA
TCCTTCCCTGACCCCAGGGGCCTCTACCACTTCTGCCTCTACTGGAACCGACATGCTGGGAGATTACATCTTCTC
TATGGCAAGCGTGACTTCTTGCTGAGTGACAAAGCCTCTAGCCTCCTCTGCTTCCAGCACCAGGAGGAGAGCCTG
GCTCAGGGCCCCCCGCTGTTAGCCACTTCTGTCACCTCCTGGTGGAGCCCTCAGAACATCAGCCTGCCCAGTGCC
GCCAGCTTCACCTTCTCCTTCCACAGTCCTCCCCACACGGCCGCTCACAATGCCTCGGTGGACATGTGCGAGCTC
AAAAGGGACCTCCAGCTGCTCAGCCAGTTCCTGAAGCATCCCCAGAAGGCCTCAAGGAGGCCCTCGGCTGCCCCC
GCCAGCCAGCAGTTGCAGAGCCTGGAGTCGAAACTGACCTCTGTGAGATTCATGGGGGACATGGTGTCCTTCGAG
GAGGACCGGATCAACGCCACGGTGTGGAAGCTCCAGCCCACAGCCGGCCTCCAGGACCTGCACATCCACTCCCGG
CAGGAGGAGGAGCAGAGCGAGATCATGGAGTACTCGGTGCTGCTGCCTCGAACACTCTTCCAGAGGACGAAAGGC
CGGAGCGGGGAGGCTGAGAAGAGACTCCTCCTGGTGGACTTCAGCAGCCAAGCCCTGTTCCAGGACAAGAATTCC
AGCCAAGTCCTGGGTGAGAAGGTCTTGGGGATTGTGGTACAGAACACCAAAGTAGCCAACCTCACGGAGCCCGTG
GTGCTCACTTTCCAGCACCAGCTACAGCCGAAGAATGTGACTCTGCAATGTGTGTTCTGGGTTGAAGACCCCACA
TTGAGCAGCCCGGGGCATTGGAGCAGTGCTGGGTGTGAGACCGTCAGGAGAGAAACCCAAACATCCTGCTTCTGC
AACCACTTGACCTACTTTGCAGTGCTGATGGTCTCCTCGGTGGAGGTGGACGCCGTGCACAAGCACTACCTGAGC
CTCCTCTCCTACGTGGGCTGTGTCGTCTCTGCCCTGGCCTGCCTTGTCACCATTGCCGCCTACCTCTGCTCCAGG
GTGCCCCTGCCGTGCAGGAGGAAACCTCGGGACTACACCATCAAGGTGCACATGAACCTGCTGCTGGCCGTCTTC
CTGCTGGACACGAGCTTCCTGCTCAGCGAGCCGGTGGCCCTGACAGGCTCTGAGGCTGGCTGCCGAGCCAGTGCC
ATCTTCCTGCACTTCTCCCTGCTCACCTGCCTTTCCTGGATGGGCCTCGAGGGGTACAACCTCTACCGACTCGTG
GTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTT
CTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAACTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAG
GGCGTCATCTACCCTTCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCAACCTGGGCCTCTTCAGC
CTGGTGTTTCTGTTCAACATGGCCATGCTAGCCACCATGGTGGTGCAGATCCTGCGGCTGCGCCCCCACACCCAA
AAGTGGTCACATGTGCTGACACTGCTGGGCCTCAGCCTGGTCCTTGGCTGCCCTGGGCCTTGATCTTCTTCTCC
TTTGCTTCTGGCACCTTCCAGCTTGTCGTCCTCTACCTTTTCAGCATCATCACCTCCTTCCAAGGCTTCCTCATC
TTCATCTGGTACTGGTCCATGCGGCTGCAGGCCCGGGGTGGCCCCTCCCCTCTGAAGAGCAACTCAGACAGCGCC
AGGCTCCCCATCAGCTCGGGCAGCACCTCGTCCAGCCGCATC**TAG**GCCTCCAGCCCACCTGCCCATGTGATGAAG
CAGAGATGCGGCCTCGTCGCACACTGCCTGTGGCCCCCGAGCCAGGCCCAGCCCCAGGCCAGTCAGCCGCAGACT
TTGGAAAGCCCAACGACCATGGAGAGATGGGCCGTTGCCATGGTGGACGGACTCCCGGGCTGGGCTTTTGAATTG
GCCTTGGGGACTACTCGGCTCTCACTCAGCTCCCACGGGACTCAGAAGTGCGCCGCCATGCTGCCTAGGGTACTG
TCCCCACATCTGTCCCAACCCAGCTGGAGGCCTGGTCTCTCCTTACAACCCCTGGGCCCAGCCCTCATTGCTGGG
GGCCAGGCCTTGGATCTTGAGGGTCTGGCACATCCTTAATCCTGTGCCCCTGCCTGGGACAGAAATGTGGCTCCA
GTTGCTCTGTCTCTCGTGGTCACCCTGAGGGCACTCTGCATCCTCTGTCATTTTAACCTCAGGTGGCACCCAGGG
CGAATGGGGCCCAGGGCAGACCTTCAGGGCCAGAGCCCTGGCGGAGGAGAGGCCCTTTGCCAGGAGCACAGCAGC
AGCTCGCCTACCTCTGAGCCCAGGCCCCCTCCCTCCCTCAGCCCCCCAGTCCTCCCTCCATCTTCCCTGGGGTTC
TCCTCCTCTCCCAGGGCCTCCTTGCTCCTTCGTTCACAGCTGGGGGTCCCCGATTCCAATGCTGTTTTTTGGGGA
GTGGTTTCCAGGAGCTGCCTGGTGTCTGCTGTAAATGTTTGTCTACTGCACAAGCCTCGGCCTGCCCCTGAGCCA
GGCTCGGTACCGATGCGTGGGCTGGGCTAGGTCCCTCTGTCCATCTGGGCCTTTGTATGAGCTGCATTGCCCTTG
CTCACCCTGACCAAGCACACGCCTCAGAGGGGCCCTCAGCCTCTCCTGAAGCCCTCTTGTGGCAAGAACTGTGGA
CCATGCCAGTCCCGTCTGGTTTCCATCCCACCACTCCAAGGACTGAGACTGACCTCCTCTGGTGACACTGGCCTA
GAGCCTGACACTCTCCTAAGGAGGTTCTCTCCAAGCCCCCAAATAGCTCCAGGCGCCCTCGGCCGCCCATCATGGT
TAATTCTGTCCAACAAACACACACGGGTAGATTGCTGGCCTGTTGTAGGTGGTAGGGACACAGATGACCGACCTG
GTCACTCCTCCTGCCAACATTCAGTCTGGTATGTGAGGCGTGCGTGAAGCAAGAACTCCTGGAGCTACAGGGACA
GGGAGCCATCATTCCTGCCTGGGAATCCTGGAAGACTTCCTGCAGGAGTCAGCGTTCAATCTTGACCTTGAAGAT
GGGAAGGATGTTCTTTTTACGTACCAATTCTTTTGTCTTTTGATATTAAAAAGAAGTACATGTTCATTGTAGAGA
ATTTGGAAACTGTAGAAGAGAATCAAGAAGAAAAATAAAAATCAGCTGTTGTAATCGCCTAGCAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 204

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921

<subunit 1 of 1, 693 aa, 1 stop

<MW: 77738, pI: 8.87, NX(S/T): 7

MTPQSLLQTTLFLLSLLFLVQGAHGRGHREDFRFCSQRNQTHRSSLHYKPTPDLRISIENSE
EALTVHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLLSDKASSLLCFQH
QEESLAQGPPLLATSVTSWWSPQNISLPSAASFTFSFHSPPHTAAHNASVDMCELKRDLQLL
SQFLKHPQKASRRPSAAPASQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTLFQRTKGRSGEAEKRLLLVDFSSQALFQDKNSSQVLGE
KVLGIVVQNTKVANLTEPVVLTFQHQLQPKNVTLQCVFWVEDPTLSSPGHWSSAGCETVRRE
TQTSCFCNHLTYFAVLMVSSVEVDAVHKHYLSLLSYVGCVVSALACLVTIAAYLCSRVPLPC
RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPVALTGSEAGCRASAIFLHFSLLTCLSWMGLE
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVDNYGPIILAVHRTPEGVIY
PSMCWIRDSLVSYITNLGLFSLVFLFNMAMLATMVVQILRLRPHTQKWSHVLTLLGLSLVLG
LPWALIFFSFASGTFQLVVLYLFSIITSFQGFLIFIWYWSMRLQARGGPSPLKSNSDSARLP
ISSGSTSSSRI

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

FIGURE 205

TGCCTGGCCTGCCTTGTCAACAATGCCGCTTACTCTGCTTCCAGGTTGCCCTGCCTTGCAGA

GGAAANCNTCGGGACTACACCNTCAAGTGCACATGAACCTGCTGCTGGCCGTCTTCCTGCTG

GACACGAGCTTCCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA

GCCAGTGCCATCTTCCTGCACTTCTCCTGCTCACCTGCCTTTCCTGGATGGGCCTCGAGGGG

TACAACCTCTACCGACTCGTGGTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAA

GCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTTCTGGTGACGCTGGTGGCCCTGGTGGATG

TGGACAACTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCCT

TCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCAACCTGGGCCTCTTCAGCCT

GGTGTTTCTGTTCAACATGG

FIGURE 206

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTCAGGTCCAGGTTTTGCTTTGA
TCCTTTTCAAAAACTGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTTTGGATGGGATTATGTGGAAACTACCCT
GCGATTCTCTGCTGCCAGAGCAGGCTCGGCGCTTCCACCCCAGTGCAGCCTTCCCCTGGCGGTGGTGAAAGAGAC
TCGGGAGTCGCTGCTTCCAAAGTGCCCGCCGTGAGTGAGCTCTCACCCCAGTCAGCCAA**ATG**AGCCTCTTCGGGC
TTCTCCTGCTGACATCTGCCCTGGCCGGCCAGAGACAGGGGACTCAGGCGGAATCCAACCTGAGTAGTAAATTCC
AGTTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTATTACTGTGTCTACTAATG
GAAGTATTCACAGCCCAAGGTTTCCTCATACTTATCCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG
AGGAAAATGTATGGATACAACTTACGTTTGATGAAAGATTTGGGCTTGAAGACCCAGAAGATGACATATGCAAGT
ATGATTTTGTAGAAGTTGAGGAACCCAGTGATGGAACTATATTAGGGCGCTGGTGTGGTTCTGGTACTGTACCAG
GAAAACAGATTTCTAAAGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCCTTCTGAACCAGGGT
TCTGCATCCACTACAACATTGTCATGCCACAATTCACAGAAGCTGTGAGTCCTTCAGTGCTACCCCCTTCAGCTT
TGCCACTGGACCTGCTTAATAATGCTATAACTGCCTTTAGTACCTTGGAAGACCTTATTCGATATCTTGAACCAG
AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTTGGCAAGGCTTTTGTTTTTGGAA
GAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACTTCT
CAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTG
GTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACC
ACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGC
ACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGA**TAG**CCGCATCACCACCAGCAGCTCTTGCCCA
GAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGCT
TCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCA
ACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTAT
TAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTC
GATACGGCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAAC
TCTAAAGCTCCATGTCCTGGGCCTAAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTTGCTCATATTCACAT
ATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAAACTTGT
GTCATGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAAAATTTCTGCCATTTAGAAGAAGAGAACTACA
TTCATGGTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTG
TTTCATTGTGTACATTTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTAAATATATCT
ATTTTTTACCAAAGGTATTTAATATTCTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCT
AAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCA
TTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAA
GACTTTTTGAAAATAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGA
AAGTAGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACAT
AAAGCACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCGTGCTGTGCTGTGCAGTAGGAACACATCCTATTTA
TTGTGATGTTGTGGTTTTATTATCTTAAACTCTGTTCCATACACTTGTATAAATACATGGATATTTTTATGTACA
GAAGTATGTCTCTTAACCAGTTCACTTATTGTACTCTGGCAATTTAAAAGAAAATCAGTAAAATATTTTGCTTGT
AAAATGCTTAATATNGTGCCTAGGTTATGTGGTGACTATTTGAATCAAAAATGTATTGAATCATCAAATAAAAGA
ATGTGGCTATTTTGGGGAGAAAATTAAAAAAAAAAAAAAAAAAAAAGGTTTAGGGATAACAGGGTAATGCGGCC

FIGURE 207

MSLFGLLLLTSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR

FPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC

GSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLL

NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNLLTEEVRLY

SCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPSKVTKKYHEVLQ

LRPKTGVRGLHKSLTDVALEHHEECDCVCRGSTGG

Signal sequence:

amino acids 1-14

FIGURE 208

CCCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATTCCATTTTGGAAGA
AGACTAAAA**ATG**GTGTTTCCAATGTGGACACTGAAGAGACAAATTCTTATCCTTTTTAACATAATCCTAATTTCC
AAACTCCTTGGGGCTAGATGGTTTCCTAAAACTCTGCCCTGTGATGTCACTCTGGATGTTCCAAAGAACCATGTG
ATCGTGGACTGCACAGACAAGCATTTGACAGAAATTCCTGGAGGTATTCCCACGAACACCACGAACCTCACCCTC
ACCATTAACCACATACCAGACATCTCCCCAGCGTCCTTTCACAGACTGGACCATCTGGTAGAGATCGATTTCAGA
TGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCAGATTAAACCCAGAAGC
TTTAGTGGACTCACTTATTTAAAATCCCTTTACCTGGATGGAAACCAGCTACTAGAGATACCGCAGGGCCTCCCG
CCTAGCTTACAGCTTCTCAGCCTTGAGGCCAACAACATCTTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCC
AACATAGAAATACTCTACCTGGGCCAAAACTGTTATTATCGAAATCCTTGTTATGTTTCATATTCAATAGAGAAA
GATGCCTTCCTAAACTTGACAAAGTTAAAAGTGCTCTCCCTGAAAGATAACAATGTCACAGCCGTCCCTACTGTT
TTGCCATCTACTTTAACAGAACTATATCTCTACAACAACATGATTGCAAAAATCCAAGAAGATGATTTTAATAAC
CTCAACCAATTACAAATTCTTGACCTAAGTGGAAATTGCCCTCGTTGTTATAATGCCCCATTTCCTTGTGCGCCG
TGTAAAAATAATTCTCCCCTACAGATCCCTGTAAATGCTTTTGATGCGCTGACAGAATTAAAAGTTTTACGTCTA
CACAGTAACTCTCTTCAGCATGTGCCCCCAAGATGGTTTAAGAACATCAACAAACTCCAGGAACTGGATCTGTCC
CAAAACTTCTTGGCCAAAGAAATTGGGGATGCTAAATTTCTGCATTTTCTCCCCAGCCTCATCCAATTGGATCTG
TCTTTCAATTTTGAACTTCAGGTCTATCGTGCATCTATGAATCTATCACAAGCATTTTCTTCACTGAAAAGCCTG
AAAATTCTGCGGATCAGAGGATATGTCTTTAAAGAGTTGAAAAGCTTTAACCTCTCGCCATTACATAATCTTCAA
AATCTTGAAGTTCTTGATCTTGGCACTAACTTTATAAAAATTGCTAACCTCAGCATGTTTAAACAATTTAAAAGA
CTGAAAGTCATAGATCTTTCAGTGAATAAAATATCACCTTCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAAT
GCCAGAACTTCTGTAGAAAGTTATGAACCCCAGGTCCTGGAACAATTACATTATTTCAGATATGATAAGTATGCA
AGGAGTTGCAGATTCAAAAACAAAGAGGCTTCTTTCATGTCTGTTAATGAAAGCTGCTACAAGTATGGGCAGACC
TTGGATCTAAGTAAAAATAGTATATTTTTTGTCAAGTCCTCTGATTTTCAGCATCTTTCTTTCCTCAAATGCCTG
AATCTGTCAGGAAATCTCATTAGCCAAACTCTTAATGGCAGTGAATTCCAACCTTTAGCAGAGCTGAGATATTTG
GACTTCTCCAACAACCGGCTTGATTTACTCCATTCAACAGCATTTGAAGAGCTTCACAAACTGGAAGTTCTGGAT
ATAAGCAGTAATAGCCATTATTTTCAATCAGAAGGAATTACTCATATGCTAAACTTTACCAAGAACCTAAAGGTT
CTGCAGAAACTGATGATGAACGACAATGACATCTCTTCCTCCACCAGCAGGACCATGGAGAGTGAGTCTCTTAGA
ACTCTGGAATTCAGAGGAAATCACTTAGATGTTTTATGGAGAGAAGGTGATAACAGATACTTACAATTATTCAAG
AATCTGCTAAAATTAGAGGAATTAGACATCTCTAAAAATTCCCTAAGTTTCTTGCCTTCTGGAGTTTTTGATGGT
ATGCCTCCAAATCTAAAGAATCTCTCTTTGGCCAAAAATGGGCTCAAATCTTTCAGTTGGAAGAAACTCCAGTGT
CTAAAGAACCTGGAAACTTTGGACCTCAGCCACAACCAACTGACCACTGTCCCTGAGAGATTATCCAACTGTTCC
AGAAGCCTCAAGAATCTGATTCTTAAGAATAATCAAATCAGGAGTCTGACGAAGTATTTTCTACAAGATGCCTTC
CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTTCCCAGAAAATGTCCTC
AACAATCTGAAGATGTTGCTTTTGCATCATAATCGGTTTCTGTGCACCTGTGATGCTGTGTGGTTTGTCTGGTGG
GTTAACCATACGGAGGTGACTATTCCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAAGGGC
CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTTCCATA
TCTGTATCTCTCTTTCTCATGGTGATGATGACAGCAAGTCACCTCTATTTCTGGGATGTGTGGTATATTTACCAT
TTCTGTAAGGCCAAGATAAAGGGGTATCAGCGTCTAATATCACCAGACTGTTGCTATGATGCTTTTATTGTGTAT
GACACTAAAGACCCAGCTGTGACCGAGTGGGTTTTGGCTGAGCTGGTGGCCAAACTGGAAGACCCAAGAGAGAAA
CATTTTAATTTATGTCTCGAGGAAAGGGACTGGTTACCAGGGCAGCCAGTTCTGGAAAACCTTTCCCAGAGCATA
CAGCTTAGCAAAAAGACAGTGTTTGTGATGACAGACAAGTATGCAAAGACTGAAAATTTTAAGATAGCATTTTAC
TTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGTGATTATCTTGATATTTCTTGAGAAGCCCTTTCAGAAG
TCCAAGTTCCTCCAGCTCCGGAAAAGGCTCTGTGGGAGTTCTGTCCTTGAGTGGCCAACAAACCCGCAAGCTCAC
CCATACTTCTGGCAGTGTCTAAAGAACGCCCTGGCCACAGACAATCATGTGGCCTATAGTCAGGTGTTCAAGGAA
ACGGTC**TAG**CCCTTCTTTGCAAAACACAACTGCCTAGTTTACCAAGGAGAGGCCTGGC

FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGG

IPTNTTNLTLTINHIPDISPASFHRLDHLVEIDFRCNCVPIPLGSKNNMCIKRLQIKPRSFS

GLTYLKSLYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLGQNCYYR

NPCYVSYSIEKDAFLNLTKLKVLSLKDNNVTAVPTVLPSTLTELYLYNNMIAKIQEDDFNNL

NQLQILDLSGNCPRCYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWF

KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSLKSL

KILRIRGYVFKELKSFNLSPLHNLQNLEVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKIS

PSGDSSEVGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQ

TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNNRLDLLH

STAFEELHKLEVLDISSNSHYFQSEGITHMLNFTKNLKVLQKLMMNDNDISSSTSRTMESES

LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNSLSFLPSGVFDGMPPNLKNLSL

AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLQ

DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLLHHNRFLCTCDAVWFVWWVNHTEVTIP

YLATDVTCVGPGAHKGQSVISLDLYTCELDLTNLILFSLSISVSLFLMVMMTASHLYFWDVW

YIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPAVTEWVLAELVAKLEDPREKHFNLCLEE

RDWLPGQPVLENLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVIILIFLE

KPFQKSKFLQLRKRLCGSSVLEWPTNPQAHPYFWQCLKNALATDNHVAYSQVFKETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 210

GGGTACCATTCTGCGCTGCTGCAAGTTACGGAATGAAAAATTAGAACAACAGAAAC**ATG**GAAAACATGTTCCTTC
AGTCGTCAATGCTGACCTGCATTTTCCTGCTAATATCTGGTTCCTGTGAGTTATGCGCCGAAGAAAATTTTTCTA
GAAGCTATCCTTGTGATGAGAAAAAGCAAAATGACTCAGTTATTGCAGAGTGCAGCAATCGTCGACTACAGGAAG
TTCCCCAAACGGTGGGCAAATATGTGACAGAACTAGACCTGTCTGATAATTTCATCACACACATAACGAATGAAT
CATTTCAAGGGCTGCAAAATCTCACTAAAATAAATCTAAACCACAACCCCAATGTACAGCACCAGAACGGAAATC
CCGGTATACAATCAAATGGCTTGAATATCACAGACGGGGCATTCCTCAACCTAAAAAACCTAAGGGAGTTACTGC
TTGAAGACAACCAGTTACCCCAAATACCCTCTGGTTTGCCAGAGTCTTTGACAGAACTTAGTCTAATTCAAAACA
ATATATACAACATAACTAAAGAGGGCATTTCAAGACTTATAAACTTGAAAAATCTCTATTTGGCCTGGAACTGCT
ATTTTAACAAAGTTTGCGAGAAAACTAACATAGAAGATGGAGTATTTGAAACGCTGACAAATTTGGAGTTGCTAT
CACTATCTTTCAATTCTCTTTCACACGTGCCACCCAAACTGCCAAGCTCCCTACGCAAACTTTTTCTGAGCAACA
CCCAGATCAAATACATTAGTGAAGAAGATTTCAAGGGATTGATAAATTTAACATTACTAGATTTAAGCGGGAACT
GTCCGAGGTGCTTCAATGCCCCATTTCCATGCGTGCCTTGTGATGGTGGTGCTTCAATTAATATAGATCGTTTTG
CTTTTCAAAACTTGACCCAACTTCGATACCTAAACCTCTCTAGCACTTCCCTCAGGAAGATTAATGCTGCCTGGT
TTAAAAATATGCCTCATCTGAAGGTGCTGGATCTTGAATTCAACTATTTAGTGGGAGAAATAGTCTCTGGGGCAT
TTTTAACGATGCTGCCCCGCTTAGAAATACTTGACTTGTCTTTTAACTATATAAAGGGGAGTTATCCACAGCATA
TTAATATTTCCAGAAACTTCTCTAAACTTTTGTCTCTACGGGCATTGCATTTAAGAGGTTATGTGTTCCAGGAAC
TCAGAGAAGATGATTTCCAGCCCCTGATGCAGCTTCCAAACTTATCGACTATCAACTTGGGTATTAATTTTATTA
AGCAAATCGATTTCAAACTTTTCCAAAATTTCTCCAATCTGGAAATTATTTACTTGTCAGAAAACAGAATATCAC
CGTTGGTAAAAGATACCCGGCAGAGTTATGCAAATAGTTCCTCTTTTCAACGTCATATCCGGAAACGACGCTCAA
CAGATTTTGAGTTTGACCCACATTCGAACTTTTATCATTTCACCCGTCCTTTAATAAAGCCACAATGTGCTGCTT
ATGGAAAAGCCTTAGATTTAAGCCTCAACAGTATTTTCTTCATTGGGCCAAACCAATTTGAAAATCTTCCTGACA
TTGCCTGTTTAAATCTGTCTGCAAATAGCAATGCTCAAGTGTTAAGTGGAACTGAATTTTCAGCCATTCCTCATG
TCAAATATTTGGATTTGACAAACAATAGACTAGACTTTGATAATGCTAGTGCTCTTACTGAATTGTCCGACTTGG
AAGTTCTAGATCTCAGCTATAATTCACACTATTTCAGAATAGCAGGCGTAACACATCATCTAGAATTTATTCAAA
ATTTCACAAATCTAAAAGTTTTAAAACTTGAGCCACAACAACATTTATACTTTAACAGATAAGTATAACCTGGAAA
GCAAGTCCCTGGTAGAATTAGTTTTCAGTGGCAATCGCCTTGACATTTTGTGGAATGATGATGACAACAGGTATA
TCTCCATTTTCAAAGGTCTCAAGAATCTGACACGTCTGGATTTATCCCTTAATAGGCTGAAGCACATCCCAAATG
AAGCATTCCTTAATTTGCCAGCGAGTCTCACTGAACTACATATAAATGATAATATGTTAAAGTTTTTTAACTGGA
CATTACTCCAGCAGTTTCCTCGTCTCGAGTTGCTTGACTTACGTGGAAACAAACTACTCTTTTTAACTGATAGCC
TATCTGACTTTACATCTTCCCTTCGGACACTGCTGCTGAGTCATAACAGGATTTCCCACCTACCCTCTGGCTTTC
TTTCTGAAGTCAGTAGTCTGAAGCACCTCGATTTAAGTTCCAATCTGCTAAAAACAATCAACAAATCCGCACTTG
AAACTAAGACCACCACCAAATTATCTATGTTGGAACTACACGGAAACCCCTTTGAATGCACCTGTGACATTGGAG
ATTTCCGAAGATGGATGGATGAACATCTGAATGTCAAAATTCCCAGACTGGTAGATGTCATTTGTGCCAGTCCTG
GGGATCAAAGAGGGAAGAGTATTGTGAGTCTGGAGCTAACAACTTGTGTTTCAGATGTCACTGCAGTGATATTAT
TTTTCTTCACGTTCTTTATCACCACCATGGTTATGTTGGCTGCCCTGGCTCACCATTTGTTTTACTGGGATGTTT
GGTTTATATATAATGTGTGTTTAGCTAAGGTAAAAGGCTACAGGTCTCTTTCCACATCCCAAACTTTCTATGATG
CTTACATTTCTTATGACACCAAAGATGCCTCTGTTACTGACTGGGTGATAAATGAGCTGCGCTACCACCTTGAAG
AGAGCCGAGACAAAAACGTTCTCCTTTGTCTAGAGGAGAGGGATTGGGACCCGGGATTGGCCATCATCGACAACC
TCATGCAGAGCATCAACCAAAGCAAGAAAACAGTATTTGTTTTAACCAAAAAATATGCAAAAAGCTGGAACTTTA
AAACAGCTTTTTACTTGGCTTTGCAGAGGCTAATGGATGAGAACATGGATGTGATTATATTTATCCTGCTGGAGC
CAGTGTTACAGCATTCTCAGTATTTGAGGCTACGGCAGCGGATCTGTAAGAGCTCCATCCTCCAGTGGCCTGACA
ACCCGAAGGCAGAAGGCTTGTTTTGGCAAACTCTGAGAAATGTGGTCTTGACTGAAAATGATTCACGGTATAACA
ATATGTATGTCGATTCCATTAAGCAATAC**TAA**CTGACGTTAAGTCATGATTTCGCGCCATAATAAAGATGCAAAG
GAATGACATTTCTGTATTAGTTATCTATTGCTATGTAACAAATTATCCCAAAACTTAGTGGTTTAAAACAACACA
TTTGCTGGCCCACAGTTTTTGAGGGTCAGGAGTCCAGGCCCAGCATAACTGGGTCCTCTGCTCAGGGTGTCTCAG
AGGCTGCAATGTAGGTGTTCACCAGAGACATAGGCATCACTGGGGTCACACTCATGTGGTTGTTTTCTGGATTCA
ATTCCTCCTGGGCTATTGGCCAAAGGCTATACTCATGTAAGCCATGCGAGCCTCTCCCACAAGGCAGCTTGCTTC
ATCAGAGCTAGCAAAAAAGAGAGGTTCTAGCAAGATGAAGTCACAATCTTTTGTAATCGAATCAAAAAAGTGAT
ᵢATCTCATCACTTTGGCCATATTCTATTTGTTAGAAGTAAACCACAGGTCCCACCAGCTCCATGGGAGTGACCACC
TCAGTCCAGGGAAAACAGCTGAAGACCAAGATGGTGAGCTCTGATTGCTTCAGTTGGTCATCAACTATTTTCCCT
TGACTGCTGTCCTGGGATGGCCTGCTATCTTGATGATAGATTGTGAATATCAGGAGGCAGGGATCACTGTGGACC
ATCTTAGCAGTTGACCTAACACATCTTCTTTTCAATATCTAAGAACTTTTGCCACTGTGACTAATGGTCCTAATA
TTAAGCTGTTGTTTATATTTATCATATATCTATGGCTACATGGTTATATTATGCTGTGGTTGCGTTCGGTTTTAT
TTACAGTTGCTTTTACAAATATTTGCTGTAACATTTGACTTCTAAGGTTTAGATGCCATTTAAGAACTGAGATGG
ATAGCTTTTAAAGCATCTTTTACTTCTTACCATTTTTTAAAAGTATGCAGCTAAATTCGAAGCTTTTGGTCTATA
TTGTTAATTGCCATTGCTGTAAATCTTAAAATGAATGAATAAAAATGTTTCATTTTACAAAAAAAAAAAAAAAAA

FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSYPCDEKKQNDSVIAECSNRRLQEVPQTVG

KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHQNGNPGIQSNGLNITDGAFLNL

KNLRELLLEDNQLPQIPSGLPESLTELSLIQNNIYNITKEGISRLINLKNLYLAWNCYFNKV

CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLINL

TLLDLSGNCPRCFNAPFPCVPCDGGASINIDRFAFQNLTQLRYLNLSSTSLRKINAAWFKNM

PHLKVLDLEFNYLVGEIVSGAFLTMLPRLEILDLSFNYIKGSYPQHINISRNFSKLLSLRAL

HLRGYVFQELREDDFQPLMQLPNLSTINLGINFIKQIDFKLFQNFSNLEIIYLSENRISPLV

KDTRQSYANSSSFQRHIRKRRSTDFEFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI

GPNQFENLPDIACLNLSANSNAQVLSGTEFSAIPHVKYLDLTNNRLDFDNASALTELSDLEV

LDLSYNSHYFRIAGVTHHLEFIQNFTNLKVLNLSHNNIYTLTDKYNLESKSLVELVFSGNRL

DILWNDDDNRYISIFKGLKNLTRLDLSLNRLKHIPNEAFLNLPASLTELHINDNMLKFFNWT

LLQQFPRLELLDLRGNKLLFLTDSLSDFTSSLRTLLLSHNRISHLPSGFLSEVSSLKHLDLS

SNLLKTINKSALETKTTTKLSMLELHGNPFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP

GDQRGKSIVSLELTTCVSDVTAVILFFFTFFITTMVMLAALAHHLFYWDVWFIYNVCLAKVK

GYRSLSTSQTFYDAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWDPGLAIID

NLMQSINQSKKTVFVLTKKYAKSWNFKTAFYLALQRLMDENMDVIIFILLEPVLQHSQYLRL

RQRICKSSILQWPDNPKAEGLFWQTLRNVVLTENDSRYNNMYVDSIKQY

Signal sequence:
amino acids 1-26

Transmembrane domain:
amino acids 826-848

FIGURE 212

CCAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT

CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA

CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC

AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC

TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGA

GCAGCTCCTGCCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAA

GGCCACCCCGCCTGGAGGCACAGGCC**ATG**AGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT

CTGGTGTTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGT

CCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTCGTGCAGCGTGTGTACCAGCCCTTCCTCA

CCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC

CGCAGCCCTGGGCTGGCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGAC

CAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGA

GCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCA

GATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCGG

CAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC

CCAAGGGAGGGCCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAA

GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGC

CCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCC

TGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCTTCCTG

GAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGACTCG**TGA**CTGCCCAGCGCCCCAGGCTG

GACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCCTGCCCAACATGCTGGGGGTCCAG

AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCC

TTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCAC

CCCTGGCTACCCCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTG

AGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAG

GCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAACGTGA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGT

CGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAAT

FIGURE 213

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC

STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR

CPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTLCVPKGGPPRVA

PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG

RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAG

GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCC

AGCAGCATCAGAGCAGCCCCTGTGGTTGGCAGCAAAGTTCAGCTTGGCTGGGCCCGCTGTGA

GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTTCTCCCTAGC

AGTGGATGAGCAACCCAACGGGGGCCCGGGGAGGGGAACTGGCCCCGAGGGAGAGGAACCCC

AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGG

CACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCCTGTCCGGGGGATGACTGATTC

TCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGCCTGGAGGCACAGGCC**ATG**AGGGGC

TCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTTGGCAGTGGGCGGCACAGAGCACGCCTA

CCGGCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTCG

TGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTAC

CGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGGGCTGGCCCCTGCCAGGCCTCGCTA

CGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATAT

GCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCA

GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTG

TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCC

TGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCG

ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCT

GGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGC

ATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGAC

TCCCTGAGCGAGCAGATTTCCTTCCTGGAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGA

CTCG**TGA**CTGCCCAGCGCTCCAGGCTGGACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATG

CCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC

AGGGCCTTCCTCCTCTTCCTCCTCCCCTTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGAT

GGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAACGGCA

TCCCAAGGCCAGGTGGACCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGAC

CCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGAC

CCCCAGCACAATAAAAATGAAACGTG

FIGURE 215

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC

STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR

CPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTLCVPKGGPPRVA

PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG

RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 216

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA

GGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAGGG

TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCAT

CAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGC

CCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGC

CTGGAGGCACAGGCC**ATG**AGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTTGGC

AGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACG

GGGACCCTGTCTCCGAGTCGTTCGTGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGAC

GGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGG

GCTGGCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTC

CTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAG

CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGA

ATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCGTCAACACCGCCGGCAGTTACTGGT

GCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGG

CCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG

GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACA

GCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCC

TTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCTTCCTGGAGGAGCAGCT

GGGGTCCTGCTCCTGCAAGAAAGACTCG**TGA**CTGCCCAGCGCCCCAGGCTGGACTGAGCCCC

TCACGCCGCCCTGCAGCCCCCATGCCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCG

GGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCCTTCCTCGGGAG

GCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACC

CCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTAC

GAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGG

CCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAACGTG

FIGURE 217

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCVNTAGSYWCQCWEGHSLSADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 218

GGTTGCCACAGCTGGTTTAGGGCCCCGACCACTGGGGCCCCCTTGTCAGGAGGAGACAGCCTCCCGGCCCGGGGAG
GACAAGTCGCTGCCACCTTTGGCTGCCGACGTGATTCCCTGGGACGGTCCGTTTCCTGCCGTCAGCTGCCGGCCG
AGTTGGGTCTCCGTGTTTCAGGCCGGCTCCCCCTTCCTGGTCTCCCTTCTCCCGCTGGGCCGGTTTATCGGGAGG
AGATTGTCTTCCAGGGCTAGCAATTGGACTTTTGATGATGTTTGACCCAGCGGCAGGAATAGCAGGCAACGTGAT
TTCAAAGCTGGGCTCAGCCTCTGTTTCTTCTCTCGTGTAATCGCAAAACCCATTTTGGAGCAGGAATTCCAATC**A**
TGTCTGTGATGGTGGTGAGAAAGAAGGTGACACGGAAATGGGAGAAACTCCCAGGCAGGAACACCTTTTGCTGTG
ATGGCCGCGTCATGATGGCCCGGCAAAAGGGCATTTTCTACCTGACCCTTTTCCTCATCCTGGGGACATGTACAC
TCTTCTTCGCCTTTGAGTGCCGCTACCTGGCTGTTCAGCTGTCTCCTGCCATCCCTGTATTTGCTGCCATGCTCT
TCCTTTTCTCCATGGCTACACTGTTGAGGACCAGCTTCAGTGACCCTGGAGTGATTCCTCGGGCGCTACCAGATG
AAGCAGCTTTCATAGAAATGGAGATAGAAGCTACCAATGGTGCGGTGCCCCAGGGCCAGCGACCACCGCCTCGTA
TCAAGAATTTCCAGATAAACAACCAGATTGTGAAACTGAAATACTGTTACACATGCAAGATCTTCCGGCCTCCCC
GGGCCTCCCATTGCAGCATCTGTGACAACTGTGTGGAGCGCTTCGACCATCACTGCCCCTGGGTGGGGAATTGTG
TTGGAAAGAGGAACTACCGCTACTTCTACCTCTTCATCCTTTCTCTCTCCCTCCTCACAATCTATGTCTTCGCCT
TCAACATCGTCTATGTGGCCCTCAAATCTTTGAAAATTGGCTTCTTGGAGACATTGAAAGAAACTCCTGGAACTG
TTCTAGAAGTCCTCATTTGCTTCTTTACACTCTGGTCCGTCGTGGGACTGACTGGATTTCATACTTTCCTCGTGG
CTCTCAACCAGACAACCAATGAAGACATCAAAGGATCATGGACAGGGAAGAATCGCGTCCAGAATCCCTACAGCC
ATGGCAATATTGTGAAGAACTGCTGTGAAGTGCTGTGTGGCCCCTTGCCCCCCAGTGTGCTGGATCGAAGGGGTA
TTTTGCCACTGGAGGAAAGTGGAAGTCGACCTCCCAGTACTCAAGAGACCAGTAGCAGCCTCTTGCCACAGAGCC
CAGCCCCCACAGAACACCTGAACTCAAATGAGATGCCGGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAG
AGCCCCCAGAGCCACCACAGGAGGCAGCTGAAGCTGAGAAG**TAG**CCTATCTATGGAAGAGACTTTTGTTTGTGTT
TAATTAGGGCTATGAGAGATTTCAGGTGAGAAGTTAAACCTGAGACAGAGAGCAAGTAAGCTGTCCCTTTTAACT
GTTTTTCTTTGGTCTTTAGTCACCCAGTTGCACACTGGCATTTTCTTGCTGCAAGCTTTTTTAAATTTCTGAACT
CAAGGCAGTGGCAGAAGATGTCAGTCACCTCTGATAACTGGAAAAATGGGTCTCTTGGGCCCTGGCACTGGTTCT
CCATGGCCTCAGCCACAGGGTCCCCTTGGACCCCCTCTCTTCCCTCCAGATCCCAGCCCTCCTGCTTGGGGTCAC
TGGTCTCATTCTGGGGCTAAAAGTTTTTGAGACTGGCTCAAATCCTCCCAAGCTGCTGCACGTGCTGAGTCCAGA
GGCAGTCACAGAGACCTCTGGCCAGGGGATCCTAACTGGGTTCTTGGGGTCTTCAGGACTGAAGAGGAGGGAGAG
TGGGGTCAGAAGATTCTCCTGGCCACCAAGTGCCAGCATTGCCCACAAATCCTTTTAGGAATGGGACAGGTACCT
TCCACTTGTTGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTGTTTTTCCTTTTGACTCCTGCTCCCATTAGGAG
CAGGAATGGCAGTAATAAAAGTCTGCACTTTGGTCATTTCTTTTCCTCAGAGGAAGCCCGAGTGCTCACTTAAAC
ACTATCCCCTCAGACTCCCTGTGTGAGGCCTGCAGAGGCCCTGAATGCACAAATGGGAAACCAAGGCACAGAGAG
GCTCTCCTCTCCTCTCCTCTCCCCCGATGTACCCTCAAAAAAAAAAAAAATGCTAACCAGTTCTTCCATTAAGCCT
CGGCTGAGTGAGGGAAAGCCCAGCACTGCTGCCCTCTCGGGTAACTCACCCTAAGGCCTCGGCCCACCTCTGGCT
ATGGTAACCACACTGGGGGCTTCCTCCAAGCCCCGCTCTTCCAGCACTTCCACCGGCAGAGTCCCAGAGCCACTT
CACCCTGGGGGTGGGCTGTGGCCCCCAGTCAGCTCTGCTCAGGACCTGCTCTATTTCAGGGAAGAAGATTTATGT
ATTATATGTGGCTATATTTCCTAGAGCACCTGTGTTTTCCTCTTTCTAAGCCAGGGTCCTGTCTGGATGACTTAT
GCGGTGGGGGAGTGTAAACCGGAACTTTTCATCTATTTGAAGGCGATTAAACTGTGTCTAATGCA

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTLFLILGTCTLFFAFECRYLAV

QLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGQRPPP

RIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCPWVGNCVGKRNYRYFYL

FILSLSLLTIYVFAFNIVYVALKSLKIGFLETLKETPGTVLEVLICFFTLWSVVGLTGFHTF

LVALNQTTNEDIKGSWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGSR

PPSTQETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMPPPEPPEPPQEAAEAEK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 220

AAAACCCTGTATTTTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTTGAATTAGGTAT

TATAGGGATGGTGGGGTTGATTTTTTNTTCCTGGAGGCTTTTGGCTTTGGACTCTCNCTTTCT

CCCACAGAGCNCTTCGACCATCACTGCCCCTGGGTGGGGAATTGTGTTGGAAAGAGGAACTA

CCGCTANTTCTACCTCTTCATCCTTTNTCTCTCCCNCCTCACAATCTATGTCTTCGCCTTCA

ACATCGT

FIGURE 221

GTTGTGTCCTTCAGCAAAACAGTGGATTTAAATCTCCTTGCACAAGCTTGAGAGCAACACAA

TCTATCAGGAAAGAAAGAAAGAAAAAAACCGAACCTGACAAAAAAGAAGAAAAAGAAGAAGA

AAAAAAATC**ATG**AAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC

GGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC

CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATT

GACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGA

CAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCG

AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC

CACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAAATTGTAGAGATTTC

TTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC

CAGAGCCTACGGTTACTTGGAGACACATCTCTCCCAAAGCGGTTGGCTTTGTGAGTGAAGAC

GAATACTTGGAAATTCAGGGCATCACCCGGGAGCAGTCAGGGACTACGAGTGCAGTGCCTC

CAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAAGGTCACCGTGAACTATCCACCATACA

TTTCAGAAGCCAAGGGTACAGGTGTCCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCC

TCAGCAGTCCCCTCAGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA

GAAAGGGGTGAAAGTGGAAAACAGACCTTTCCTCTCAAAACTCATCTTCTTCAATGTCTCTG

AACATGACTATGGGAACTACACTTGCGTGGCCTCCAACAAGCTGGGCCACACCAATGCCAGC

ATCATGCTATTTGGTCCAGGCGCCGTCAGCGAGGTGAGCAACGGCACGTCGAGGAGGGCAGG

CTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGCACCTGCTTCTCAAATTT**TGA**TGTGAGTGCC

ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCACCAACACAACAGCAATGGCAACAC

CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA

AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGGAAAAGAGTTTTAAAAAAGAAATTGAA

AATTGCCTTGCAGATATTTAGGTACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGC

ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA

GGGCTCAGCCTCTCTGCCCACAGAGTGCCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA

AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG

GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAGAGCAAAA

AAAAA

FIGURE 222

MKTIQPKMHNSISWAIFTGLAALCLFQGVPVRSGDATFPKAMDNVTVRQGESATLRCTIDNR

VTRVAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPK

TSRVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDEYL

EIQGITREQSGDYECSASNDVAAPVVRRVKVTVNYPPYISEAKGTGVPVGQKGTLQCEASAV

PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNYTCVASNKLGHTNASIML

FGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF

Signal peptide:

amino acids 1-28

FIGURE 223

GAAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTC

ACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTT

CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTA

TTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT

GACAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCAT

CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA

ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAAATTGTAGAGATT

TCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAG

ACCAGAG

FIGURE 224

ATGGCTGGTGACGGCGGGGCCGGGCAGGGGACCGGGGCCGCGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGA
ATCACCGCCTGGCCCGACTCCACC**ATG**AACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAG
AAGGGGACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCT
GCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCATCCCACAGCACCTGCCTTACA
GAGGCCTGCATTCGAGTGGCTGGAAAAATCCTGGAGTCCCTGGACCGAGGGGTGAGCCCCTGTGAGGACTTTTAC
CAGTTCTCCTGTGGGGGCTGGATTCGGAGGAACCCCCTGCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGC
CTCTGGGACCAAAACCAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAG
CAGAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCCAGCCACTGAGA
GACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAGGACAACTTTATGGAGGTGTTGAAG
GCAGTAGCAGGGACCTACAGGGCCACCCCATTCTTCACCGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGC
AATGTTATCCAGGTGGACCAGTCTGGGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAG
AAAGTGCTCACTGCCTATCTGGATTACATGGAGGAACTGGGGATGCTGCTGGGTGGGCGGCCCACCTCCACGAGG
GAGCAGATGCAGCAGGTGCTGGAGTTGGAGATACAGCTGGCCAACATCACAGTGCCCCAGGACCAGCGGCGCGAC
GAGGAGAAGATCTACCACAAGATGAGCATTTCGGAGCTGCAGGCTCTGGCGCCCTCCATGGACTGGCTTGAGTTC
CTGTCTTTCTTGCTGTCACCATTGGAGTTGAGTGACTCTGAGCCTGTGGTGGTGTATGGGATGGATTATTTGCAG
CAGGTGTCAGAGCTCATCAACCGCACGGAACCAAGCATCCTGAACAATTACCTGATCTGGAACCTGGTGCAAAAG
ACAACCTCAAGCCTGGACCGACGCTTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCCTCTATGGCACTAAGAAG
TCCTGTGTGCCGAGGTGGCAGACCTGCATCTCCAACACGGATGACGCCCTTGGCTTTGCTTTGGGGTCACTCTTC
GTGAAGGCCACGTTTGACCGGCAAAGCAAAGAAATTGCAGAGGGGATGATCAGCGAAATCCGGACCGCATTTGAG
GAGGCCCTGGGACAGCTGGTTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAGGAGAAAGCAGATGCCATCTAT
GATATGATTGGTTTCCCAGACTTTATCCTGGAGCCCAAAGAGCTGGATGATGTTTATGACGGGTACGAAATTTCT
GAAGATTCTTTCTTCCAAAACATGTTGAATTTGTACAACTTCTCTGCCAAGGTTATGGCTGACCAGCTCCGCAAG
CCTCCCAGCCGAGACCAGTGGAGCATGACCCCCCAGACAGTGAATGCCTACTACCTTCCAACTAAGAATGAGATC
GTCTTCCCCGCTGGCATCCTGCAGGCCCCCTTCTATGCCCGCAACCACCCCAAGGCCCTGAACTTCGGTGGCATC
GGTGTGGTCATGGGCCATGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTG
CGGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAGTACAATCAA
TACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACATTACTGACAACGGGGGGCTGAAG
GCTGCCTACAATGCTTACAAAGCATGGCTGAAAGACATGGGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACC
AACCACCAGCTCTTCTTCGTGGGATTTGCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGG
CTGGTGACCGACCCCCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAACTCCCGTGACTTCCTGCGG
CACTTCGGCTGCCCTGTCGGCTCCCCCATGAACCAGGGCAGCTGTGTGAGGTGTGG**TAG**ACCTGGATCAGGGGA
GAAATGGCCAGCTGTCACCAGACCTGGGGCAGCTCTCCTGACAAAGCTGTTTGCTCTTGGGTTGGGAGGAAGCAA
ATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCACAGGTGACATGAGTACAGACCCTCCTCAATCACCACATTG
TGCCTCTGCTTTGGGGGTGCCCCTGCCTCCAGCAGAGCCCCCACCCATTCACTGTGACATCTTTCCGTGTCACCCT
GCCTGGAAGAGGTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCC

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDPSH

STCLTEACIRVAGKILESLDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQNQA

ILKHLLENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNITGPWDQDN

FMEVLKAVAGTYRATPFFTVYISADSKSSNSNVIQVDQSGLFLPSRDYYLNRTANEKVLTAY

LDYMEELGMLLGGRPTSTREQMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAP

SMDWLEFLSFLLSPLELSDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL

DRRFESAQEKLLETLYGTKKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI

SEIRTAFEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF

FQNMLNLYNFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLPTKNEIVFPAGILQAPFYARNH

PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV

NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCS

VRTPESSHEGLVTDPHSPARFRVLGTLSNSRDFLRHFGCPVGSPMNPGQLCEVW

Type II Transmembrane domain:
amino acids 32-57

FIGURE 226

GCCCGGCCCTCCGCCCTCCGCACTCCCGCCTCCCTCCCTCCGCCCGCTCCCGCGCCCTCCTCCCTCCCTCCTCCC
CAGCTGTCCCGTTCGCGTCATGCCGAGCCTCCCGGCCCCGCCGGCCCCGCTGCTGCTCCTCGGGCTGCTGCTGCT
CGGCTCCCGGCCGGCCCGCGGCGCCGGCCCAGAGCCCCCCGTGCTGCCCATCCGTTCTGAGAAGGAGCCGCTGCC
CGTTCGGGGAGCGGCAGGTAGGTGGGCGCCCGGGGGAGGCGCGGGGCGGGGAGTCGGGCTCGGGGCGAGTCAGCGC
CAGCCCGGAGGGGGCGCGGGGCGCAGGTGGCTCGGCGCGGCGGGCGGCCCGGAGGGTGGGCGGGGGCAGAAGGGC
GCGGTGCCTGGGACCCGGGACCCGCGGGCAGCCCCCGGGGCGGCACACGGCGCGGAGCTGGGCAGCGGCCTCCAGC
CAAGCCCGTCCCCGCAGGCTGCACCTTCGGCGGGAAGGTCTATGCCTTGGACGAGACGTGGCACCCGGACCTAGG
GGAGCCATTCGGGGTGATGCGCTGCGTGCTGTGCGCCTGCGAGGCGCAGTGGGGTCGCCGTACCAGGGGCCCTGG
CAGGGTCAGCTGCAAGAACATCAAACCAGAGTGCCCAACCCCGGCCTGTGGGCAGCCGCGCCAGCTGCCGGGACA
CTGCTGCCAGACCTGCCCCCAGGACTTCGTGGCGCTGCTGACAGGGCCGAGGTCGCAGGCGGTGGCACGAGCCCG
AGTCTCGCTGCTGCGCTCTAGCCTCCGCTTCTCTATCTCCTACAGGCGGCTGGACCGCCCTACCAGGATCCGCTT
CTCAGACTCCAATGGCAGTGTCCTGTTTGAGCACCCTGCAGCCCCCACCCAAGATGGCCTGGTCTGTGGGGTGTG
GCGGGCAGTGCCTCGGTTGTCTCTGCGGCTCCTTAGGGCAGAACAGCTGCATGTGGCACTTGTGACACTCACTCA
CCCTTCAGGGGAGGTCTGGGGGCCTCTCATCCGGCACCGGGCCCTGTCCCCAGAGACCTTCAGTGCCATCCTGAC
TCTAGAAGGCCCCCACCAGCAGGGCGTAGGGGGCATCACCCTGCTCACTCTCAGTGACACAGAGGACTCCTTGCA
TTTTTTGCTGCTCTTCCGAGGCCTTGCAGGACTAACCCAGGTTCCCTTGAGGCTCCAGATTCTACACCAGGGGCA
GCTACTGCGAGAACTTCAGGCCAATGTCTCAGCCCAGGAACCAGGCTTTGCTGAGGTGCTGCCCAACCTGACAGT
CCAGGAGATGGACTGGCTGGTGCTGGGGGAGCTGCAGATGGCCCTGGAGTGGGCAGGCAGGCCAGGGCTGCGCAT
CAGTGGACACATTGCTGCCAGGAAGAGCTGCGACGTCCTGCAAAGTGTCCTTTGTGGGGCTAATGCCCTGATCCC
AGTCCAAACGGGTGCTGCCGGCTCAGCCAGCCTCACTCTGCTAGGAAATGGCNCCCTGATCCTCCAGGTGCAATT
GGTAGGGACAACCAGTGAGGTGGTGGCCATGACACTGGAAACCAAGCCTCAGCGGAGGGATCAGCCCACTGTCCT
GTGCCACATGGCTGGCCTATCCTCCCCTGCCCCAGGCCGTGGGTATCTGCCCTGGGCTGGGGTGCCCGAGGGGC
TCATATGCTGCTGCAGAATGAGCTCTTCCTGAACGTGGGCACCAAGGACTTCCCAGACGGAGAGCTTCGGGGGCA
ACGTGGCTGCCCTGCCCTACTGTGGGGCATAGCGCCCGCCCTGCCCGTGCCCCTAGCAGGAGCCCTGGTGCTACC
CCCTGTGAAGAGCCAAGCAGCAGGGCACGCCTGGCTTTCCTTGGATACCCACTGTCACCTGCACTATGAAGTGCT
GCTGGCTGGGCTTGGTGGCTCAGAACAAGGCACTGTCACTGCCCACCTCCTTGGGCCTCCTGGAACGCCAGGGCC
TCGGCGGCTGCTGAAGGGATTCTATGGCTCAGAGGCCCAGGGTGTGGTGAAGGACCTGGAGCCGGAACTGCTGCG
GCACCTGGCAAAAGGCATGGCTTCCCTGATGATCACCACCAAGGTAGCCCCAGAGGGGAGCTCCGAGGGCAGCCT
CTCCTCCCAGGTGCACATAGCCAACCAATGTGAGGTTGGCGGACTGCGCCTGGAGGCGGCCGGGGCCGAGGGGGT
GCGGGCGCTGGGGGCTCCGGATACAGCCTCTGCTGCGCCGCCTGTGGTGCCTGGTCTCCCGGCCCTAGCGCCCGC
CAAACCTGGTGGTCCTGGGCGGCCCCGAGACCCCAACACATGCTTCTTCGAGGGGCAGCAGCGCCCCCACGGGGC
TCGCTGGGCGCCCAACTACGACCCGCTCTGCTCACTCTGCACCTGCCAGAGACGAACGGTGATCTGTGACCCGGT
GGTGTGCCCACCGCCCAGCTGCCCACACCCGGTGCAGGCTCCCGACCAGTGCTGCCCTGTTTGCCCTGGCTGCTA
TTTTGATGGTGACCGGAGCTGGCGGGCAGCGGGTACGCGGTGGCACCCCGTTGTGCCCCCCTTTGGCTTAATTAA
GTGTGCTGTCTGCACCTGCAAGCAGGGGGGCACTGGAGAGGTGCACTGTGAGAAGGTGCAGTGTCCCCGGCTGGC
CTGTGCCCAGCCTGTGCGTGTCAACCCCACCGACTGCTGCAAACAGTGTCCAGGTGAGGCCCACCCCCAGCTGGG
GGACCCCATGCAGGCTGATGGGCCCCGGGGCTGCCGTTTTGCTGGGCAGTGGTTCCCAGAGAGTCAGAGCTGGCA
CCCCTCAGTGCCCCCGTTTGGAGAGATGAGCTGTATCACCTGCAGATGTGGGGTAAGTGGGGAGCAGAGGCTTGT
GTGAGGTGGGTACTGGGAGCCTGGTCTGGAGTAGGGAGACCTTCCCAGGGGAGGTCCCTGAAGAAGCTGAAGGTCA
CTGTGTCCCAGTGCCTCTGGGGGACACTCAGTGTCTGCTCTGTCTTGTACCAGGCAGGGGTGCCTCACTGTGAGC
GGGATGACTGTTCACTGCCACTGTCCTGTGGCTCGGGGAAGGAGAGTCGATGCTGTTCCCGCTGCACGGCCCACC
GGCGGCGTAAGTGAGGGAGTCCAGGGTCAGCAGCTGTGAGTGGAGGGCTCACCTGCCTGTGGGACTCCTGATCAG
GGAAGGGAGCACTCACTGTGTGCAGGAACAGTGCAGCCTGCCTCACAAGTGCCATTCCAATCCACCCTCACAGCA
ACCTGGTGGAATTGTTATTTATGACCTTTTCTTTACAAATGAGATTTCTGAAGCTCAGAGAAATTAAGCAACGAG
ATGAAGGTCACCCAGCTGTGTGCACTGACCTGTTTAGAAAATACTGGCCTTTCTGGGACCAAGGCAGGGATGCTT
TGCCCTGCCCTCTATGCCTCTCTGTGCCTCTCCACTCCCTCTCCCCTCCTCCAACATTCCCTCCCTTCTGTCTCC
AGCAGCCCCAGAGACCAGAACTGATCCAGAGCTGGAGAAAGAAGCCGAAGGCTCTTAGGGAGCAGCCAGAGGGCC
AAGTGACCAAGAGGATGGGGCCTGAGCTGGGGAAGGGGTGGCATCGAGGACCTTCTTGCATTCTCCTGTGGGAAG
CCCAGTGCCTTTGCTCCTCTGTCCTGCCTCTACTCCCACCCCCACTACCTCTGGGAACCACAGCTCCACAAGGGG
GAGAGGCAGCTGGGCCAGACCGAGGTCACAGCCACTCCAAGTCCTGCCCTGCCACCCTCGGCCTCTGTCCTGGAA
GCCCCACCCCTTTCTTCCTGTACATAATGTCACTGGCTTGTTGGGATTTTTAATTTATCTTCACTCAGCACCAAG
GGCCCCGGACACTCCACTCCTGCTGCCCCTGAGCTGAGCAGAGTCATTATTGGAGAGTTTTGTATTTATTAAAAC
ATTTCTTTTTCAGTCTTTGGGCATGAGGTTGGCTCTTTGTGGCCAGGAACCTGAGTGGGCCTGGTGGAGAAGGG
GCNGAGAGTAGGAGGTGAGAGAGAGGAGCTCTGACACTTGGGGAGCTGAAAGAGACCTGGAGAGGCAGAGGATAG
CGTGGCNNTTGGCTGGCATNCCTGGGTTCCGCAGAGGGGCTGGGGATGGTTCTTGAGATGGTCTAGAGACTCAAG
AATTTAGGGAAGTAGAAGCAGGATTTTGACTCAAGTTTAGTTTCCCACATCGCTGGCCTGTTTGCTGACTTCATG
TTTGAAGTTGCTCCAGAGAGAGAATCAAAGGTGTCACCAGCCCCTCTCTCCCTCCTTCCCTTCCCTTCCCTTTCT
TTCCCTCCCCTCCCCTCCCCTCCCCTCCCCTCC

FIGURE 227

GGCCGAGCGGGGGTGCTGCGCGGCGGCCGTGATGGCTGGTGACGGCGGGGCCGGGCAGGGGA

CCGGGGCCGCGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCCTGGCCCGAC

TCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG

GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTAC

TGCTGGCTGCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCA

TCCCACAGCACCTGCCTTACAGAGGCCTGCATTCGAGTGGCTGGAAAAATCCTGGAGTCCCT

GGACCGAGGGGTGAGCCCCTGTGAGGACTTTTACCAGTTCTCCTGTGGGGGCTGGATTCGGA

GGAACCCCCTGCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAAC

CAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAGCA

GAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC

AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAG

GACAACTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCCATTCTTCAC

CGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG

GGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTAAGGAAC

ATCTTCCGAACCCCCATCCCTACCCCTGGCTGAGCTGGGCTGATCCCTGTTGACTTTTCCCT

TTGCCAAGGGTCAGAGCAGGGAAGGTGAGCCTATCCTGTCACCTAGTGAACAAACTGCCCCT

CCTTTCTTTCTTCTTTTCTTCCTCCCTCCCTCCCTTTCTTCCCCTTTTCCTTCCTTCCTTCC

TCTTATTCTTCTAGTAGGTTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGGAATTCG

GAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTCAGTGTCGATGGGTTCATGGACCT

AGATAGGCTGATAACAAAGCTCACAAGAGGGTCCTGAGGATTCAGGAGAGACTTATGGAGCC

AGCAAAGTCTTCCTGAAGAGATTGCATTTGAGCCAGGTCCTGTAG

FIGURE 228

ATGCCTACTACCTTCCAACTAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCCC
TTCTATGCCCGCAACCACCCCAAGGCCCTGAACTTCGGTGGCATCGGTGTGGTCATGGGCCA
TGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCGGC
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCACACGGCCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACAT
TGCTGACAACGGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATTT
GCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGGCTGGTGACCGACCC
CCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACTCTCTCCAACTCCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTCGGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGGCCAGCTGTCACCAGACCTGGGGCAGCTCTCCTGACAAAGCTGT
TTGCTCTTGGGTTGGGAGGAAGCAAATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCACA
GGTGACATGAGTACAGACCCTCCTCAATCACCACATTGTGCCTCTGCTTTGGGGGTGCCCCT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCCGTGTCACCCTGCCTGGAAGAG
GTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCCTCTTCTGTCCCCAGGCTCACT
CAGCCTGGCGGCCATGGGGCCTGCCGTGCCTGCCCCACTGTGACCCACAGGCCTGGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCGCACTTAGGGGTGGACTCAGCTCTGTC
TGGCTCACCCTCACGGGCTACCCCCACCTCACCCTGTGCTCCTTGTGCCACTGCTCCCAGTG
CTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAGCCCAAGGGCCTCTGAAAGCCTCCTGC
TGCCCACTGTTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGGTACTGGTTCCTGT
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTTATTTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCT
TGGCCCTTATAGGACC

FIGURE 229

```
CCCACGCGTCCGAGCCGCCCGAGAATTAGACACACTCCGGACGCGGCCAAAAGCAACCGAGA
GGAGGGGAGGCAAAAACACCGAAAAACAAAAAGAGAGAAACAACACCCAACAACTGGGGTGG
GGGGAAGAAAGAAAGAAAAGAAACCCACCCACCCACCAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAATCCTGTGGCGCGCCGCCTGGTTCCCGGGAAGACTCGCCAGCACCAGGGGG
TGGGGGAGTGCGAGCTGAAAGCTGCTGGAGAGTGAGCAGCCCTAGCAGGGATGGAC**ATG**ATG
CTGTTGGTGCAGGGTGCTTGTTGCTCGAACCAGTGGCTGGCGGCGGTGCTCCTCAGCCTGTG
CTGCCTGCTACCCTCCTGCCTCCCGGCTGGACAGAGTGTGGACTTCCCCTGGGCGGCCGTGG
ACAACATGATGGTCAGAAAAGGGGACACGGCGGTGCTTAGGTGTTATTTGGAAGATGGAGCT
TCAAAGGGTGCCTGGCTGAACCGGTCAAGTATTATTTTTGCGGGAGGTGATAAGTGGTCAGT
GGATCCTCGAGTTTCAATTTCAACATTGAATAAAAGGGACTACAGCCTCCAGATACAGAATG
TAGATGTGACAGATGATGGCCCATACACGTGTTCTGTTCAGACTCAACATACACCCAGAACA
ATGCAGGTGCATCTAACTGTGCAAGTTCCTCCTAAGATATATGACATCTCAAATGATATGAC
CGTCAATGAAGGAACCAACGTCACTCTTACTTGTTTGGCCACTGGGAAACCAGAGCCTTCCA
TTTCTTGGCGACACATCTCCCCATCAGCAAAACCATTTGAAAATGGACAATATTTGGACATT
TATGGAATTACAAGGGACCAGGCTGGGGAATATGAATGCAGTGCGGAAAATGCTGTGTCATT
CCCAGATGTGAGGAAAGTAAAAGTTGTTGTCAACTTTGCTCCTACTATTCAGGAAATTAAAT
CTGGCACCGTGACCCCCGGACGCAGTGGCCTGATAAGATGTGAAGGTGCAGGTGTGCCGCCT
CCAGCCTTTGAATGGTACAAAGGAGAGAAGAAGCTCTTCAATGGCCAACAAGGAATTATTAT
TCAAAATTTTAGCACAAGATCCATTCTCACTGTTACCAACGTGACACAGGAGCACTTCGGCA
ATTATACCTGTGTGGCTGCCAACAAGCTAGGCACAACCAATGCGAGCCTGCCTCTTAACCCT
CCAAGTACAGCCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTTTCTCCTGCTGGTACCT
TGTGTTGACACTGTCCTCTTTCACCAGCATATTCTACCTGAAGAATGCCATTCTACAA**TAA**A
TTCAAAGACCCATAAAAGGCTTTTAAGGATTCTCTGAAAGTGCTGATGGCTGGATCCAATCT
GGTACAGTTTGTTAAAAGCAGCGTGGGATATAATCAGCAGTGCTTACATGGGGATGATCGCC
TTCTGTAGAATTGCTCATTATGTAAATACTTTAATTCTACTCTTTTTTGATTAGCTACATTA
CCTTGTGAAGCAGTACACATTGTCCTTTTTTTAAGACGTGAAAGCTCTGAAATTACTTTTAG
AGGATATTAATTGTGATTTCATGTTTGTAATCTACAACTTTTCAAAAGCATTCAGTCATGGT
CTGCTAGGTTGCAGGCTGTAGTTTACAAAAACGAATATTGCAGTGAATATGTGATTCTTTAA
GGCTGCAATACAAGCATTCAGTTCCCTGTTTCAATAAGAGTCAATCCACATTTACAAAGATG
CATTTTTTTCTTTTTTGATAAAAAAGCAAATAATATTGCCTTCAGATTATTTCTTCAAAATA
TAACACATATCTAGATTTTTCTGCTTGCATGATATTCAGGTTTCAGGAATGAGCCTTGTAAT
ATAACTGGCTGTGCAGCTCTGCTTCTCTTTCCTGTAAGTTCAGCATGGGTGTGCCTTCATAC
AATAATATTTTTCTCTTTGTCTCCAACTAATATAAAATGTTTTGCTAAATCTTACAATTTGA
AAGTAAAAATAAACCAGAGTGATCAAGTTAAACCATACACTATCTCTAAGTAACGAAGGAGC
TATTGGACTGTAAAAATCTCTTCCTGCACTGACAATGGGGTTTGAGAATTTTGCCCCACACT
AACTCAGTTCTTGTGATGAGAGACAATTTAATAACAGTATAGTAAATATACCATATGATTTC
TTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGAAATCATTCCCTTTAAAATGACAGCACA
GTCCACTCAAAGGATTGCCTAGCAATACAGCATCTTTTCCTTTCACTAGTCCAAGCCAAAAA
TTTTAAGATGATTTGTCAGAAAGGGCACAAAGTCCTATCACCTAATATTACAAGAGTTGGTA
AGCGCTCATCATTAATTTTATTTTGTGGCAGGTATTATGACAGTCGACCTGGAGGGTATGGA
TATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACCAGGGTGGTTATGACCGCTACTC
AGGAGGAAATTACAGAGACAATTATGACAACTGAAATGAGACATGCACATAATATAGATACA
CAAGGAATAATTTCTGATCCAGGATCGTCCTTCCAAATGGCTGTATTTATAAAGGTTTTTGG
AGCTGCACTGAAGCATCTTATTTTATAGTATATCAACCTTTTGTTTTTAAATTGACCTGCCA
AGGTAGCTGAAGACCTTTTAGACAGTTCCATCTTTTTTTTTTTAAATTTTTTTCTGCCTATTTAA
AGACAAATTATGGGACGTTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 230

MMLLVQGACCSNQWLAAVLLSLCCLLPSCLPAGQSVDFPWAAVDNMMVRKGDTAVLRCYLED
GASKGAWLNRSSIIFAGGDKWSVDPRVSISTLNKRDYSLQIQNVDVTDDGPYTCSVQTQHTP
RTMQVHLTVQVPPKIYDISNDMTVNEGTNVTLTCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRDQAGEYECSAENAVSFPDVRKVKVVVNFAPTIQEIKSGTVTPGRSGLIRCEGAGV
PPPAFEWYKGEKKLFNGQQGIIIQNFSTRSILTVTNVTQEHFGNYTCVAANKLGTTNASLPL
NPPSTAQYGITGSADVLFSCWYLVLTLSSFTSIFYLKNAILQ

Important features of the protein:
Signal peptide:
amino acids 1-31

Transmembrane domain:
amino acids 326-345

N-glycosylation sites.
amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.
amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.
amino acids 178-186

N-myristoylation sites.
amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297, 302-308, 319-325

Myelin P0 protein:
amino acids 92-121

FIGURE 231

AGTGGTTCGATGGGAAGGATCTTTCTCCAAGTGGTTCCTCTTGAGGGGAGCATTTCTGCTGG

CTCCAGGACTTTGGCCATCTATAAAGCTTGGCA**ATG**AGAAATAAGAAAATTCTCAAGGAGGA

CGAGCTCTTGAGTGAGACCCAACAAGCTGCTTTTCACCAAATTGCAATGGAGCCTTTCGAAA

TCAATGTTCCAAAGCCCAAGAGGAGAAATGGGGTGAACTTCTCCCTAGCTGTGGTGGTCATC

TACCTGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGC

GCGGCTCCGGGTCCTGGAGATGTATTTCCTCAATGACACTCTGGCGGCTGAGGACAGCCCGT

CCTTCTCCTTGCTGCAGTCAGCACACCCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTG

CAAGTCCTGCAGGCCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGT

AGACAACTTCACTCAGAACCCAGGGATGTTCAGAATCAAAGGTGAACAAGGCGCCCCAGGTC

TTCAAGGTCACAAGGGGGCCATGGGCATGCCTGGTGCCCCTGGCCCGCCGGGACCACCTGCT

GAGAAGGGAGCCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCCTCGGGACCCCAAGG

CCCACCGGGAGTCAAGGGAGAGGCGGGCCTCCAAGGACCCCAGGGTGCTCCAGGGAAGCAAG

GAGCCACTGGCACCCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATT

GGCCCAAAAGGGGAAACTGGAACTAAGGGAGAGAAAGGAGACCTGGGTCTCCCAGGAAGCAA

AGGGGACAGGGGCATGAAAGGAGATGCAGGGGTCATGGGGCCTCCTGGAGCCCAGGGGAGTA

AAGGTGACTTCGGGAGGCCAGGCCCACCAGGTTTGGCTGGTTTTCCTGGAGCTAAAGGAGAT

CAAGGACAACCTGGACTGCAGGGTGTTCCGGGCCCTCCTGGTGCAGTGGGACACCCAGGTGC

CAAGGGTGAGCCTGGCAGTGCTGGCTCCCCTGGGCGAGCAGGACTTCCAGGGAGCCCCGGGA

GTCCAGGAGCCACAGGCCTGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGACAGCAAGGA

AGAAAAGGAGAATCAGGAGTTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGCCCAGG

GCTGGCAGGTCCCAAGGGAGCCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAAAG

GATCTTCTGGGGAGCAAGGAGTAAAGGGAGAAAAAGGTGAAAGAGGTGAAAACTCAGTGTCC

GTCAGGATTGTCGGCAGTAGTAACCGAGGCCGGGCTGAAGTTTACTACAGTGGTACCTGGGG

GACAATTTGCGATGACGAGTGGCAAAATTCTGATGCCATTGTCTTCTGCCGCATGCTGGGTT

ACTCCAAAGGAAGGGCCCTGTACAAAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAAT

GTTCAGTGTCGGGGCACGGAGAGTACCCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCA

TGACTGCAGCCACGAGGAGGACGCAGGCGTGGAGTGCAGCGTC**TGA**CCCGGAAACCCTTTCA

CTTCTCTGCTCCCGAGGTGTCCTCGGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGT

TCCCTGGGGACAACTGAGCAGCCTCTGGAGAGGGGCCATTAATAAAGCTCAACATCATTGA

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886

><subunit 1 of 1, 520 aa, 1 stop

><MW: 52658, pI: 9.16, NX(S/T): 3

MRNKKILKEDELLSETQQAAFHQIAMEPFEINVPKPKRRNGVNFSLAVVVIYLILLTAGAGL
LVVQVLNLQARLRVLEMYFLNDTLAAEDSPSFSLLQSAHPGEHLAQGASRLQVLQAQLTWVR
VSHEHLLQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRD
GATGPSGPQGPPGVKGEAGLQGPQGAPGKQGATGTPGPQGEKGSKGDGGLIGPKGETGTKGE
KGDLGLPGSKGDRGMKGDAGVMGPPGAQGSKGDFGRPGPPGLAGFPGAKGDQGQPGLQGVPG
PPGAVGHPGAKGEPGSAGSPGRAGLPGSPGSPGATGLKGSKGDTGLQGQQGRKGESGVPGPA
GVKGEQGSPGLAGPKGAPGQAGQKGDQGVKGSSGEQGVKGEKGERGENSVSVRIVGSSNRGR
AEVYYSGTWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVGAGTGQIWLDNVQCRGTESTLW
SCTKNSWGHHDCSHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352, 400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

C1q domain proteins.

amino acids 151-184, 301-334, 316-349

FIGURE 233

CCCACGCGTCCGAAGGCAGACAAAGGTTCATTTGTAAAGAAGCTCCTTCCAGCACCTCCTCT
CTTCTCCTTTTGCCCAAACTCACCCAGTGAGTGTGAGCATTTAAGAAGCATCCTCTGCCAAG
ACCAAAAGGAAAGAAGAAAAAGGGCCAAAAGCCAAA**ATG**AAACTGATGGTACTTGTTTTCAC
CATTGGGCTAACTTTGCTGCTAGGAGTTCAAGCCATGCCTGCAAATCGCCTCTCTTGCTACA
GAAAGATACTAAAAGATCACAACTGTCACAACCTTCCGGAAGGAGTAGCTGACCTGACACAG
ATTGATGTCAATGTCCAGGATCATTTCTGGGATGGGAAGGGATGTGAGATGATCTGTTACTG
CAACTTCAGCGAATTGCTCTGCTGCCCAAAAGACGTTTTCTTTGGACCAAAGATCTCTTTCG
TGATTCCTTGCAACAATCAA**TGA**GAATCTTCATGTATTCTGGAGAACACCATTCCTGATTTC
CCACAAACTGCACTACATCAGTATAACTGCATTTCTAGTTTCTATATAGTGCAATAGAGCAT
AGATTCTATAAATTCTTACTTGTCTAAGACAAGTAAATCTGTGTTAAACAAGTAGTAATAAA
AGTTAATTCAATCTAAAAAAAAAAAA

FIGURE 234

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758
<subunit 1 of 1, 98 aa, 1 stop
<MW: 11081, pI: 6.68, NX(S/T): 1          .
MKLMVLVFTIGLTLLLGVQAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDG
KGCEMICYCNFSELLCCPKDVFFGPKISFVIPCNNQ
```

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

FIGURE 235

CCCACGCGTCCGCGGACGCGTGGGCTGGACCCCAGGTCTGGAGCGAATTCCAGCCTGCAGGG
CTGATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTTACCCCGCCGTGGTGGTTGGAGGGC
GCGCAGTAGAGCAGCAGCACAGGCGCGGGTCCCGGGAGGCCGGCTCTGCTCGCGCCGAG**ATG**
TGGAATCTCCTTCACGAAACCGACTCGGCTGTGGCCACCGCGCGCCGCCCGCGCTGGCTGTG
CGCTGGGGCGCTGGTGCTGGCGGGTGGCTTCTTTCTCCTCGGCTTCCTCTTCGGGTGGTTTA
TAAAATCCTCCAATGAAGCTACTAACATTACTCCAAAGCATAATATGAAAGCATTTTTGGAT
GAATTGAAAGCTGAGAACATCAAGAAGTTCTTACATAATTTTACACAGATACCACATTTAGC
AGGAACAGAACAAAACTTTCAGCTTGCAAAGCAAATTCAATCCCAGTGGAAAGAATTTGGCC
TGGATTCTGTTGAGCTAGCTCATTATGATGTCCTGTTGTCCTACCCAAATAAGACTCATCCC
AACTACATCTCAATAATTAATGAAGATGGAAATGAGATTTTCAACACATCATTATTTGAACC
ACCTCCTCCAGGATATGAAAATGTTTCGGATATTGTACCACCTTTCAGTGCTTTCTCTCCTC
AAGGAATGCCAGAGGGCGATCTAGTGTATGTTAACTATGCACGAACTGAAGACTTCTTTAAA
TTGGAACGGGACATGAAAATCAATTGCTCTGGGAAAATTGTAATTGCCAGATATGGGAAAGT
TTTCAGAGGAAATAAGGTTAAAAATGCCCAGCTGGCAGGGGCCAAAGGAGTCATTCTCTACT
CCGACCCTGCTGACTACTTTGCTCCTGGGGTGAAGTCCTATCCAGACGGTTGGAATCTTCCT
GGAGGTGGTGTCCAGCGTGGAAATATCCTAAATCTGAATGGTGCAGGAGACCCTCTCACACC
AGGTTACCCAGCAAATGAATATGCTTATAGGCGTGGAATTGCAGAGGCTGTTGGTCTTCCAA
GTATTCCTGTTCATCCAATTGGATACTATGATGCACAGAAGCTCCTAGAAAAAATGGGTGGC
TCAGCACCACCAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCCTACAATGTTGGACCTGG
CTTTACTGGAAACTTTTCTACACAAAAAGTCAAGATGCACATCCACTCTACCAATGAAGTGA
CGAGAATTTACAATGTGATAGGTACTCTCAGAGGAGCAGTGGAACCAGACAGATATGTCATT
CTGGGAGGTCACCGGGACTCATGGGTGTTTGGTGGTATTGACCCTCAGAGTGGAGCAGCTGT
TGTTCATGAAATTGTGAGGAGCTTTGGAACACTGAAAAAGGAAGGGTGGAGACCTAGAAGAA
CAATTTTGTTTGCAAGCTGGGATGCAGAAGAATTTGGTCTTCTTGGTTCTACTGAGTGGGCA
GAGGAGAATTCAAGACTCCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTAT
AGAAGGAAACTACACTCTGAGAGTTGATTGTACACCGCTGATGTACAGCTTGGTACACAACC
TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTTGAAGGCAAATCTCTTTATGAAAGTTGG
ACTAAAAAAAGTCCTTCCCCAGAGTTCAGTGGCATGCCCAGGATAAGCAAATTGGGATCTGG
AAATGATTTTGAGGTGTTCTTCCAACGACTTGGAATTGCTTCAGGCAGAGCACGGTATACTA
AAAATTGGGAAACAAACAAATTCAGCGGCTATCCACTGTATCACAGTGTCTATGAAACATAT
GAGTTGGTGGAAAAGTTTTATGATCCAATGTTTAAATATCACCTCACTGTGGCCCAGGTTCG
AGGAGGGATGGTGTTTGAGCTAGCCAATTCCATAGTGCTCCCTTTTGATTGTCGAGATTATG
CTGTAGTTTTAAGAAAGTATGCTGACAAAATCTACAGTATTTCTATGAAACATCCACAGGAA
ATGAAGACATACAGTGTATCATTTGATTCACTTTTTTCTGCAGTAAAGAATTTTACAGAAAT
TGCTTCCAAGTTCAGTGAGAGACTCCAGGACTTTGACAAAAGCAACCCAATAGTATTAAGAA
TGATGAATGATCAACTCATGTTTCTGGAAAGAGCATTTATTGATCCATTAGGGTTACCAGAC
AGGCCTTTTTATAGGCATGTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTC
ATTCCCAGGAATTTATGATGCTCTGTTTGATATTGAAAGCAAAGTGGACCCTTCCAAGGCCT
GGGGAGAAGTGAAGAGACAGATTTATGTTGCAGCCTTCACAGTGCAGGCAGCTGCAGAGACT
TTGAGTGAAGTAGCC**TAA**GAGGATTTTTTAGAGAATCCGTATTGAATTTGTGTGGTATGTCA
CTCAGAAGAATCGTAATGGGTATATTGATAAATTTTAAAATTGGTATATTTGAAATAAAGT
TGAATATTATATATAA

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756

><subunit 1 of 1, 750 aa, 1 stop

><MW: 84305, pI: 6.93, NX(S/T): 10

MWNLLHETDSAVATARRPRWLCAGALVLAGGFFLLGFLFGWFIKSSNEATNITPKHNMKAFL

DELKAENIKKFLHNFTQIPHLAGTEQNFQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTH

PNYISIINEDGNEIFNTSLFEPPPPGYENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFF

KLERDMKINCSGKIVIARYGKVFRGNKVKNAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL

PGGGVQRGNILNLNGAGDPLTPGYPANEYAYRRGIAEAVGLPSIPVHPIGYYDAQKLLEKMG

GSAPPDSSWRGSLKVPYNVGPGFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDRYV

ILGGHRDSWVFGGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW

AEENSRLLQERGVAYINADSSIEGNYTLRVDCTPLMYSLVHNLTKELKSPDEGFEGKSLYES

WTKKSPSPEFSGMPRISKLGSGNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYET

YELVEKFYDPMFKYHLTVAQVRGGMVFELANSIVLPFDCRDYAVVLRKYADKIYSISMKHPQ

EMKTYSVSFDSLFSAVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAFIDPLGLP

DRPFYRHVIYAPSSHNKYAGESFPGIYDALFDIESKVDPSKAWGEVKRQIYVAAFTVQAAAE

TLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713